

INTRODUZIONE ALLA RADIOSITÀ

Questo modello per la luce diffusa, introdotto in [C.M.Goral, K.E.Torrance, D.P.Greenberg e B.Bataille, *Modeling the Interaction of Light between Diffuse Surfaces*, SIGGRAPH 84, in *Computer Graphics* 18(3), luglio 1984, 213-222], sposta l'attenzione dallo studio delle riflessioni dirette ad una analisi accurata dell'effetto cumulativo delle interriflessioni tra gli oggetti della scena.

I primi algoritmi in questo campo si occupavano della radiazione del calore dentro motori ed altri ambienti chiusi, e da questo ambito furono poi trasportati allo studio della diffusione della luce. Si calcola quanta parte dell'energia emessa o riflessa da ogni superficie raggiunge ogni altra superficie, venendone ulteriormente riflessa o assorbita. Una parte della porzione riflessa raggiunge nuovamente la superficie originale, e quindi ne modifica la luminosità. Questo porta ad una successione di correzioni successive, che si traducono nella risoluzione con metodi iterativi di un appropriato sistema lineare. La soluzione dà la luminosità di ogni superficie della scena in conseguenza dell'illuminazione che la raggiunge direttamente dalle sorgenti di luce ed indirettamente dalla diffusione da parte delle altre superficie. La potenza luminosa totale di una superficie (emessa e/o riflessa) si chiama radiosità.

Il metodo di calcolo suddivide le superficie in pezzi più piccoli di luminosità uniforme, chiamati elementi (*patch*), e risolve per iterazione il sistema lineare di cui sopra per trovare la luminosità dei vertici di ciascun elemento generata dalla luce diretta e da quella diffusa dalle altre superficie. Il procedimento iterativo aggiunge a ciascun elemento la luminosità diffusa dalle successive. Quindi nei primi passi iterativi la luminosità globale risulta ridotta, perché si perde la luminosità dovuta ai contributi degli elementi ancora non considerate. Per rendere più rapida la convergenza dell'iterazione, viene quindi eseguito un procedimento di correzione della luminosità globale per recuperare la perdita iniziale.

L' EQUAZIONE DELLA RADIOSITÀ

Le superficie considerate nel metodo di radiosità possono emettere luce propria, a differenza di quelle considerate nel ray tracing. In particolare, le sorgenti di luce in questo metodo diventano superficie (il metodo è ideale, in particolare, per trattare sorgenti di luce estese), ma ciascuna superficie, oltre ad emettere luce propria, diffonde quella che le proviene dalle altre.

Quindi ciascuna superficie della scena è modellata come una sorgente luminosa estesa, avente un'area finita. Suddividiamo l'ambiente in un numero finito di sottosuperficie sufficientemente piccole, che chiameremo elementi, che emettono e riflettono luce in modo uniforme sull'intera area. Se consideriamo ogni elemento come una sorgente di luce che, oltre ad emettere la propria luce, diffonde quella proveniente dalle altre sorgenti (ma non la riflette specularmente: è una superficie opaca), allora, per ogni $i = 1, \dots, n$, abbiamo la seguente equazione:

$$b_i = e_i + \rho_i \sum_{j=1, \dots, n} b_j F_{ji} A_j / A_i$$

Qui b_i e b_j sono le radiosità - cioè le potenze totali per unità di area (= tassi di energia emessa per unità di tempo e di area) - uscenti dagli elementi i e j , A_i ed A_j sono le aree degli elementi i e j , e_i la potenza totale per unità di area emessa dall'elemento i , ρ_i la riflettività dell'elemento i e F_{ji} il fattore di forma dell'elemento i rispetto all'elemento j , che misura quale percentuale della potenza emessa dall'elemento j arriva all'elemento i . Il fattore di forma misura quindi quanto è grande l'angolo solido coperto dalla porzione visibile dell'elemento i quando si guarda la scena dal centro dell'elemento j (la frazione dell'angolo solido complessivo coperta dalla parte visibile dell'elemento i); sarebbe più accurato fare la media degli angoli solidi coperti dall'elemento i al variare del punto di osservazione sull'elemento j , ma questo porterebbe ad una mole di calcoli eccessiva. Il j -simo addendo della sommatoria rappresenta l'energia per unità di tempo e di area che l'elemento i riceve da quello j . Questa energia, moltiplicata per il coefficiente di riflettività dell'elemento i , è la frazione proveniente dall'elemento j della energia per unità di tempo e di area emessa dall'elemento i .

Riassumendo, questa equazione indica semplicemente la relazione esistente tra l'energia emessa per unità di tempo dall'elemento i e l'energia per unità di tempo entrante in questo elemento, vista come somma della componente di luce propria e quella di luce diffusa proveniente dagli altri elementi.

E' importante notare che il 100% dell'energia emessa da ciascun elemento deve arrivare agli altri elementi (considerando anche lo sfondo come uno o più elementi). In altre parole si ha la seguente equazione della conservazione dell'energia (ovvero dell'angolo solido totale):

$$\sum_j F_{ij} = 1$$

Come accennato più sopra, quest'ultima quantità viene determinata moltiplicando per il coefficiente di riflessione dell'elemento i la somma della luce incidente proveniente dagli altri elementi, cioè la somma al variare di j della luce che lascia un'unità di area dell'elemento j e che raggiunge l'elemento i .

Come vedremo, vale la seguente relazione di reciprocità tra i fattori di forma (*reversibilità dei percorsi ottici*):

$$A_i F_{ij} = A_j F_{ji}$$

Essenzialmente questa relazione esprime, in termini della luce inviata da un elemento all'altra, la reversibilità dei percorsi ottici: se un punto di un elemento è visibile da un punto dell'altro, allora è vero anche il viceversa.

Pertanto possiamo riscrivere l'equazione precedente così:

$$b_i = e_i + \rho_i \sum_{j=1, \dots, n} b_j F_{ij} \quad (0)$$

Riordinando i termini e considerando tutti i possibili valori di i arriviamo al seguente sistema lineare:

$$\begin{pmatrix} 1 - \rho_1 F_{1-1} & -\rho_1 F_{1-2} & \cdots & -\rho_1 F_{1-n} \\ -\rho_2 F_{2-1} & 1 - \rho_2 F_{2-2} & \cdots & -\rho_2 F_{2-n} \\ \cdots & \cdots & \cdots & \cdots \\ -\rho_n F_{n-1} & -\rho_n F_{n-2} & \cdots & 1 - \rho_n F_{n-n} \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{pmatrix}$$

Il sistema è quindi il seguente:

$$Mb = e$$

dove **b** ed **e** sono rispettivamente i vettori delle radiosità e delle potenze emesse, ed **M** è la matrice

$$M_{ij} = \delta_{ij} - \rho_i F_{ij}$$

(come sempre, il simbolo di Kronecker δ_{ij} vale 1 se $i=j$ e 0 altrimenti). E' importante rendersi conto per i termini diagonali possiamo supporre

$$M_{ii} > 0$$

dal momento che

$$\rho_i < 1$$

(la radiosità serve a studiare scene diffuse, non perfettamente speculari!), e

$$F_{ii} < 1$$

a meno che l'elemento i -simo non sia l'unico elemento della scena incluso lo sfondo, nel qual caso la scena è banale: l'unico elemento ha illuminazione costante, e non occorrono metodi numerici per il rendering.

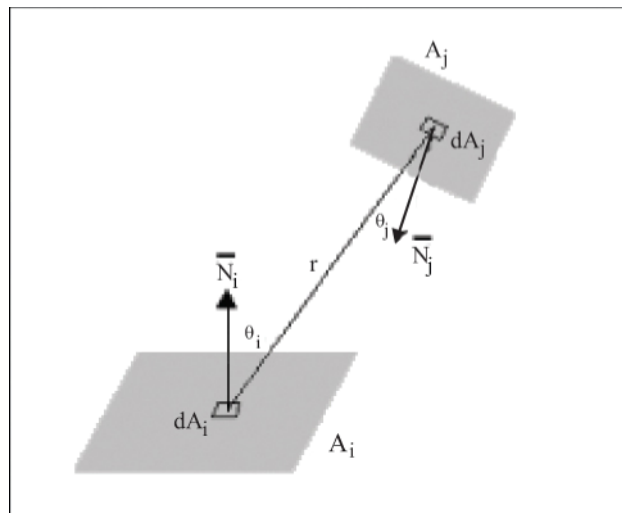
Questo sistema di equazioni, risolvibile usando il metodo iterativo di [Gauss-Seidel](#), porta ad ottenere i valori di radiosità per ogni elemento, i quali, calcolati in funzione della lunghezza d'onda, possono essere interpretati come la distribuzione spettrale dell'intensità dell'elemento nell'immagine, e sono utilizzabili nell'algoritmo di determinazione delle superfici visibili.

Nell'articolo [M.F.Cohen e D.P.Greenberg, [The Hemi-Cube: A Radiosity Solution for Complex Environments](#), SIGGRAPH 85, in *Computer Graphics* 19(3), luglio 1985, pagg. 21-40] viene sviluppato un procedimento per determinare la radiosità di un vertice della scena poligonale, utilizzabile poi nello shading interpolato (secondo il metodo di [Gouraud](#): qui non si usa il metodo di Phong perché stiamo studiando un problema di luce solo diffusa, e quindi i coefficienti di riflessione sono nulli). Il procedimento è il seguente: ad un vertice interno alla superficie viene assegnata la media delle radiosità degli elementi che condividono quel vertice. Invece, se il vertice appartiene ad un lato, la radiosità di quel vertice viene determinata in due passi. Prima si trova quella del vertice Interno più vicino al vertice dato e quelle degli elementi che condividono il vertice dato. Poi la media di questi elementi viene posta uguale alla media tra il vertice interno più vicino e la radiosità del vertice dato, che è la nostra incognita. Così si arriva ad una equazione lineare, risolvendo la quale otteniamo la quantità cercata.

CALCOLO DEI FATTORI DI FORMA

L'operazione più importante e computazionalmente più gravosa del metodo di radiosità è il calcolo dei fattori di forma, i quali misurano quale frazione dell'energia emessa da un elemento ne raggiunge un'altra. Sono quindi una misura geometrica della visibilità di una superficie da un'altra.

Nell'articolo [M.F.Cohen e D.P.Greenberg, *The Hemi-Cube: A Radiosity Solution for Complex Environments*, SIGGRAPH 85, in *Computer Graphics* 19(3), luglio 1985, pagg. 21-40] viene sviluppato un algoritmo a [precisione di immagine](#) per la determinazione delle superficie visibili per approssimare i fattori di forma. Date due elementi, la geometria che porta a determinare il fattore di forma da un elemento infinitesimo di area dA_i della prima ad un elemento infinitesimo dA_j della seconda (che chiamiamo F_{dij}) è illustrata nella figura seguente. Occorre moltiplicare le aree per i due rispettivi coseni per proiettarle perpendicolarmente al segmento che le congiunge (in generale sono disposte obliquamente rispetto ad esso, ed in tal modo coprono un angolo solido ridotto di un fattore pari a questo coseno).



Qui θ_i è l'angolo che il raggio forma con la normale alla superficie A_i , θ_j è l'angolo che il raggio forma con la normale alla superficie A_j , H_{ij} assume i valori 0 o 1 a seconda che dA_j sia visibile o meno da dA_i , e r è la lunghezza del raggio. Partendo da questo risultato possiamo ottenere il fattore di forma da un elemento infinitesimo di area dA_i all'intero elemento j -esimo di area A_j integrando su tutta l'area dell'elemento j -esimo.

In tal modo otteniamo che F_{dij} è dato da

$$F_{di-j} = \int_{A_j} \frac{\cos \theta_i \cos \theta_j}{\pi r^2} H_{ij} dA_j$$

Ora, integrando sull'elemento i-esimo e prendendo la media integrale (cioè dividendo l'integrale per l'area A_i), otteniamo il fattore di forma dall'elemento i all'elemento j:

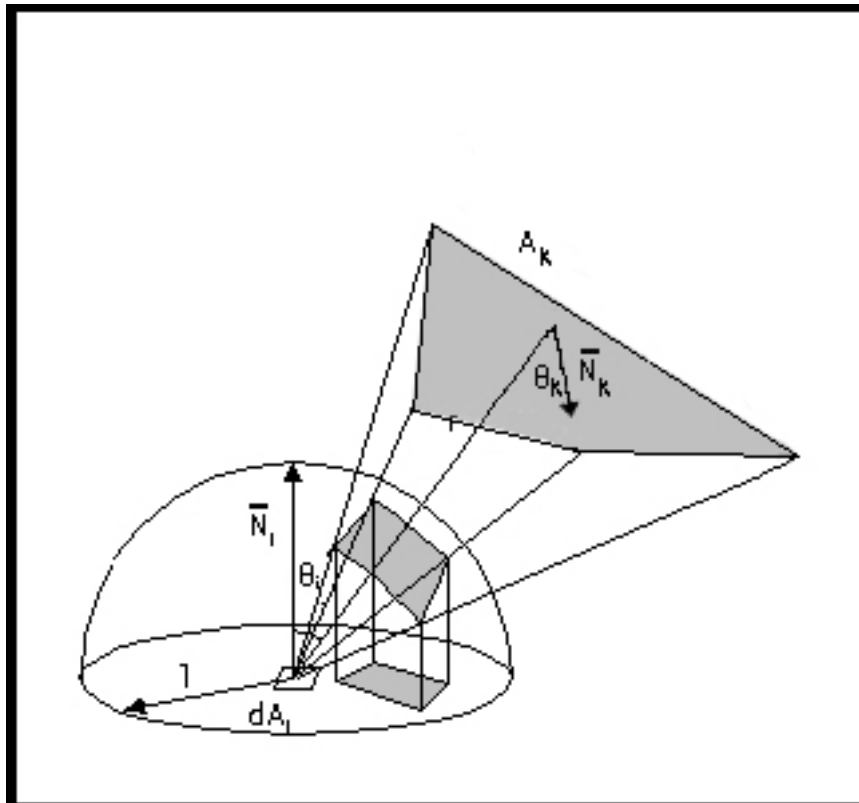
$$F_{i-j} = \frac{1}{A_i} \int_{A_i} \int_{A_j} \frac{\cos \theta_i \cos \theta_j}{\pi r^2} H_{ij} dA_j dA_i$$

Si osservi che da questa formula segue la relazione di reciprocità accennata [precedentemente](#):

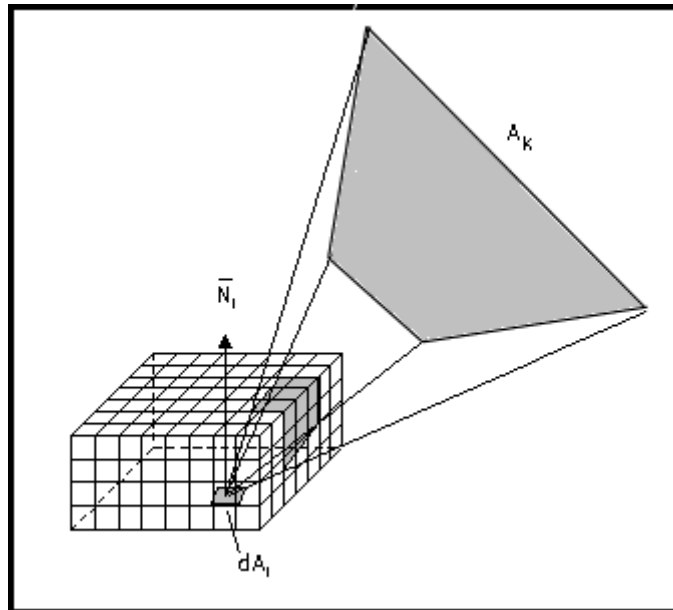
$$A_i F_{ij} = A_j F_{ji}$$

Se assumiamo che il punto centrale di un elemento rappresenti con buona precisione la luminosità di tutti gli altri suoi punti, possiamo approssimare F_{ij} con F_{dij} , calcolato per dA_i nel centro dall'elemento i-esimo.

Sono disponibili metodi numerici più semplici per calcolare F_{dij} . Nel libro di *R.Siegel e J.Howell, Thermal Radiation Heat Transfer, 2nd edition, Hemisphere, Washington, DC, 1981*) si attribuisce a Nusselt l'aver notato che la quantità F_{ik} si possa calcolare mediante due proiezioni successive: innanzitutto bisogna proiettare quelle parti di A_k (elemento k-esimo) che sono visibili dal pezzo dA_i dell'elemento A_i (elemento i-esimo) su una semisfera unitaria, centrata in dA_i : si osservi che il fattore di forma di questa proiezione (ossia l'angolo solido da essa coperto) coincide con quello dell'elemento A_k . Successivamente quest'area sulla superficie della semisfera viene proiettata ortograficamente sul disco unitario che costituisce la base della semisfera. La percentuale dell'area di questo disco coperta dalla superficie proiettata è il fattore di forma (la proiezione sulla sfera unitaria è data dalla moltiplicazione per $\cos \theta_k / r^2$, la proiezione sul disco unitario dalla moltiplicazione per $\cos \theta_i$).



Nell'articolo [M.F.Cohen e D.P.Greenberg, [The Hemi-Cube: A Radiosity Solution for Complex Environments](#), SIGGRAPH 85, in *Computer Graphics* 19(3), luglio 1985, pagg. 21-40] viene sviluppato un algoritmo a [precisione di immagine](#) (infatti basato su una decomposizione in una griglia di pixel, come lo [z-buffer](#)) nel quale la proiezione avviene sulla parte superiore di un semicubo, centrato in dA_i . Ogni faccia viene suddivisa con una griglia regolare. Tutte gli altri elementi vengono proiettate verso il centro del cubo e quindi si calcola quali celle la proiezione copre sulle cinque facce superiori. Esattamente come nel caso dell'emisfero, il fattore di forma dell'elemento originale A_k coincide con quello della sua proiezione sull'emicubo, e quindi tanto vale rimpiazzare la prima con la seconda (per la quale i vettori normali sono i vettori diretti perpendicolarmente alla faccia dell'emicubo su cui cade la proiezione). Per ciascuna cella di queste facce si determina quale sia l'elemento A_k più vicina (nella direzione di visuale dal centro del cubo) mediante l'algoritmo di z-buffer. L'elemento più vicina, per una data cella, è l'unico elemento che riflette energia luminosa su dA_i attraverso quella cella, poiché le altre sono schermate dalla prima.



Ad ogni cella viene assegnato un fattore di forma precalcolato del tipo

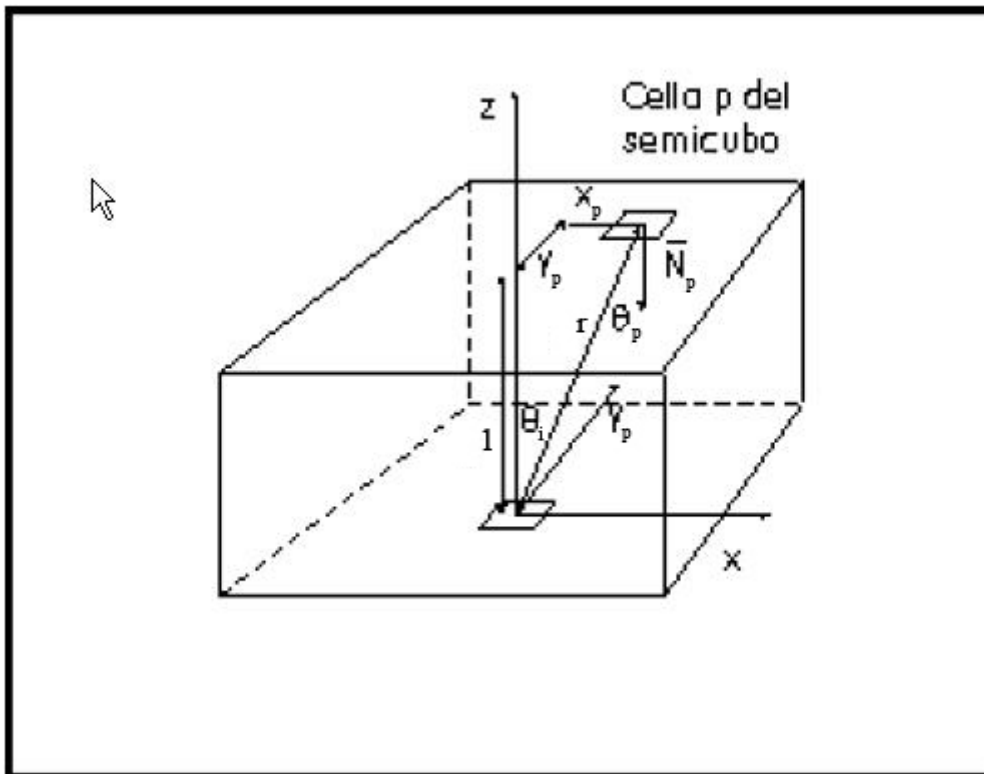
$$\Delta F_p = \frac{\cos \theta_i \cos \theta_p}{\pi r^2} \Delta A$$

dove θ_p è l'angolo tra la normale alla cella p e il vettore r che congiunge il centro di dA_i (cioè il centro del cubo) col centro di p , e ΔA è l'area della cella. In un sistema di riferimento per il semicubo con l'origine nel suo centro, per le celle della faccia superiore otteniamo:

$$r = \sqrt{1+x_p^2+y_p^2}$$

$$\cos \theta_i = \cos \theta_p = \frac{1}{r}$$

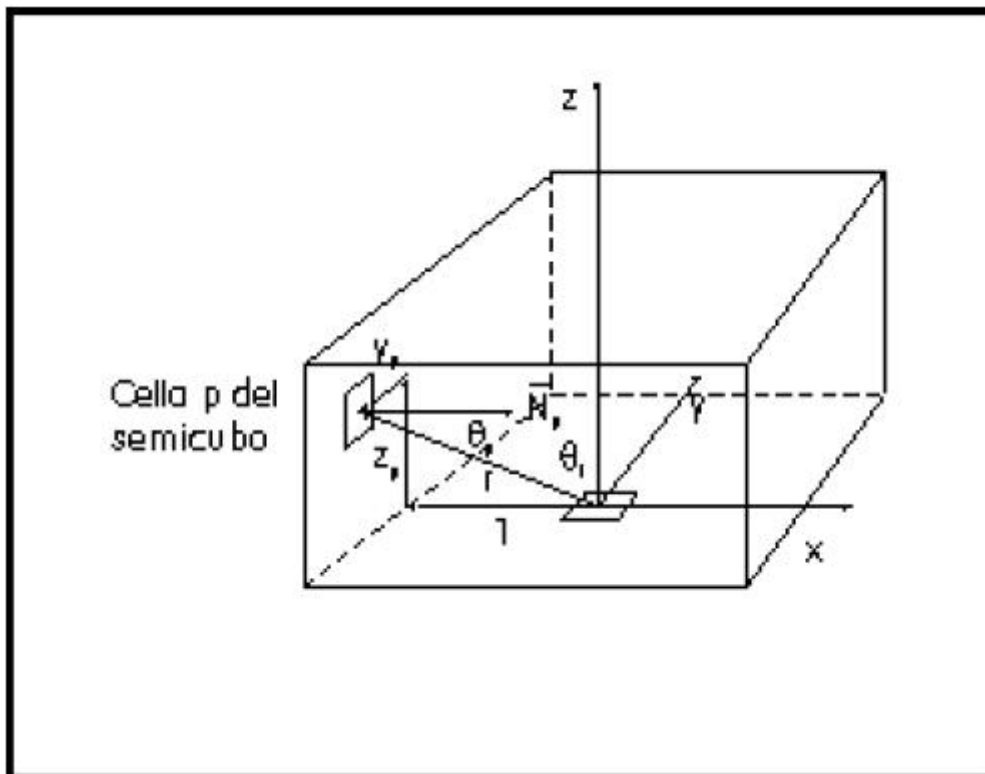
dove (x_p, y_p) sono le coordinate del centro della cella p . La geometria è illustrata nella seguente figura:



Ora l'equazione precedente, per una cella appartenente alla faccia superiore, ci dà:

$$\Delta F_p = \frac{1}{\pi(1+x_p^2+y_p^2)^2} \Delta A$$

Invece, la geometria delle celle in una faccia laterale del semicubo che è perpendicolare, diciamo, all'asse è illustrata nella figura seguente:



Analogamente, per queste celle si ha:

$$\Delta F_p = \frac{z_p}{\pi(1+x_p^2+y_p^2)^2} \Delta A$$

Qui stiamo considerando le facce laterali, ed quindi per due di esse x_p vale rispettivamente 1 o -1, e per le altre due y_p vale rispettivamente 1 o -1.

Ora il fattore di forma F_{di-k} tra l'elemento i-esimo e l'elemento k-esimo può essere approssimato con la somma di tutti i fattori di forma ΔF_p associati ad ogni cella p del semicubo centrato sull'elemento i-esimo coperta dalle proiezioni dell'elemento k-esimo. Si noti che, poiché questi metodi usano la suddivisione in celle delle facce e lo [z-buffer](#), essi sono algoritmi a [precisione di immagine](#), e quindi soggetti al fenomeno dell'[aliasing](#).

ESERCIZI SUL CALCOLO ANALITICO DEI FATTORI DI FORMA

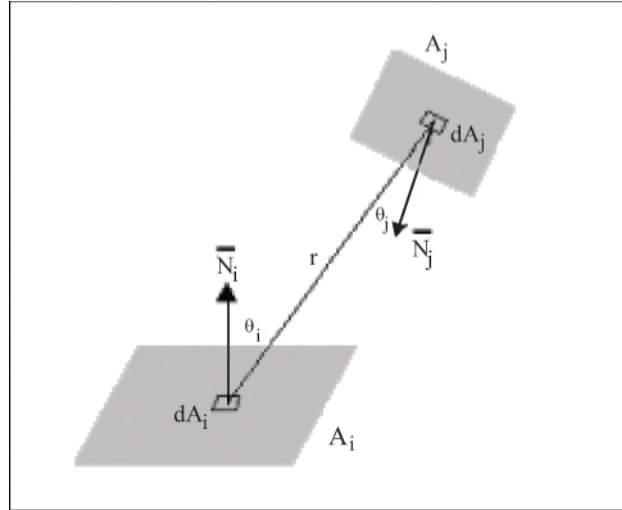
Esercizio 1. Una scena consiste di un cono retto di angolo al vertice pari a $\pi/4$, ed altezza h . Quindi la base del cono è un disco, che chiamiamo B . Supponiamo che B sia un elemento: quanto vale il suo fattore di forma visto dal vertice del cono, considerando il vertice del cono come il centro di un elemento infinitesimo di area disposto su un piano parallelo alla base? (Suggerimento: si usi il metodo di Nusselt).

Svolgimento. Per prima cosa proiettiamo il disco sulla superficie sferica di raggio 1, con proiezione radiale verso il centro di questa sfera, che è il punto da cui si vuol calcolare il fattore di forma. Per inciso, si osservi che il fattore di forma misura solo l'angolo solido sotteso, e quindi sarebbe invariato se la base del cono, invece di un disco piano, fosse una calotta sferica, ad esempio quella della sfera di raggio h , o anche di raggio 1: la prima fase del procedimento di Nusselt, la proiezione radiale, non è altro che questa semplice osservazione.

A questo punto abbiamo ottenuto una calotta C di ampiezza angolare $\pi/4$, disposta, diciamo, intorno al polo Nord ($x=y=0$, $z=1$) della superficie sferica di raggio 1 con centro l'origine. Il metodo di Nusselt consiste nell'osservare che il fattore di forma richiesto è esattamente la percentuale dell'area del disco di base della sfera (ossia il suo disco equatoriale), che vale π , data dall'area della proiezione verticale D della calotta C sul piano di base $z=0$. La deviazione angolare massima della calotta rispetto alla direzione verticale è $\pi/8$. Quindi la calotta si proietta sul disco D nel piano $z=0$ il cui raggio è il raggio della sfera per $\sin(\pi/8)$, ossia $\sin(\pi/8)$. L'area di D è $\pi \sin^2(\pi/8)$, e quindi il fattore di forma vale $\pi \sin^2(\pi/8)/\pi = \sin^2(\pi/8)$.

Esercizio 2. Per il cono dell'esercizio precedente, ricalcolare il fattore di forma della sua base B visto dal vertice del cono tramite la formula integrale che lo definisce.

Svolgimento. La formula che dà il fattore di forma è $F = \int_C \frac{\cos\theta_i \cos\theta_j}{\pi h^2} H_{ij} dC$, dove gli angoli θ_i e θ_j sono come nella figura già vista precedentemente, che riportiamo qui,



ed il fattore di occlusione H_{ij} vale 1 se il punto j è visibile dal vertice da cui vogliamo calcolare il fattore di forma (nella figura questo vertice lo chiamiamo i). Osserviamo che il fattore di occlusione vale sempre 1 perché non ci sono altri elementi che occludono la base del cono quando vista dal vertice (il cono è convesso!). Abbiamo già osservato che il fattore di forma F rimane uguale se si rimpiazza il disco di base B con la calotta sferica C di raggio h incernierata sul cono. Inoltre sui punti di C l'angolo θ_j vale sempre 0, perché la calotta sferica ha per versore normale esattamente il versore radiale. Quindi chiamiamo più semplicemente θ l'altro angolo θ_i , e notiamo che la deviazione angolare (azimuth) θ verifica $0 \leq \theta \leq \pi/8 \leq \pi/8$. Allora l'integrale diventa

$$F = \int_C \frac{\cos\theta}{\pi h^2} dC = \int_0^{2\pi} \int_0^{\pi/8} \frac{\cos\theta}{\pi h^2} h^2 \sin\theta d\theta d\varphi = \int_0^{2\pi} \sin 2\theta d\theta = \frac{1}{2} \int_0^{\pi/4} \sin\theta d\theta$$

$$= \frac{1}{2} (1 - \cos \pi/4) = \frac{1}{2} (1 - \cos^2 \pi/8 + \sin^2 \pi/8) = \sin^2 \pi/8$$

Qui il fattore $\sin \theta$ dentro l'integrale di superficie è il fattore di dilatazione delle aree della trasformazione σ da coordinate sferiche θ, φ (angoli di Eulero) a coordinate cartesiane x, y, z , a suo tempo calcolato:

$$\left\| \frac{\partial \sigma}{\partial \theta} \times \frac{\partial \sigma}{\partial \varphi} \right\| = r \sin \theta$$

Se invece si desidera svolgere il calcolo direttamente per il disco piano B , basta osservare che la normale al disco B è il versore verticale (diretto verso il basso, se il vertice del cono è all'origine ed il disco B ad altezza $h > 0$). Perciò ora $\theta_i = \theta_j$: chiamiamo questo angolo θ . Il raggio ρ_B di B è proporzionale al raggio ρ_E della calotta sferica E che B sottende. **Attenzione:** nella parte precedente di questo esercizio la calotta non era E ma C , ossia non era quella sottesa da B , bensì quella tangente a B al suo centro, che ha raggio $h < \rho_E$, ed infatti tale che

$$h/\rho_E = \cos(\pi/8) \quad (1)$$

Quindi ora il denominatore nell'integrando del fattore di forma è πr^2 invece di πh^2 : questa volta r non è costante ma dipende da x e y . Nell'Esercizio 1 abbiamo visto che $\rho_B = \rho_E \sin(\pi/8)$. Eliminiamo ρ_E sostituendo (1) in questa identità: troviamo la relazione, peraltro geometricamente evidente, $\rho_B = h \operatorname{tg}(\pi/8)$. Per la stessa ragione, un punto di B che dista ρ dal centro di B e r dall'origine ha una deviazione angolare θ rispetto alla verticale tale che $\rho = h \operatorname{tg} \theta = r \sin \theta$, e $h = r \cos \theta$ (come in (1)). Pertanto, ponendo $u = \rho/h$, otteniamo $u = \operatorname{tg} \theta$, $\cos^2 \theta = h^2/r^2$, e, per il teorema di Pitagora, $r^2 = \rho^2 + h^2$. Ne segue:

$$\begin{aligned} F &= \int_B \frac{\cos^2 \theta}{\pi r^2} d\sigma = \int_0^{2\pi} \int_0^{\operatorname{tg} \pi/8} \frac{h^2}{\pi r^4} \rho d\rho d\varphi = 2 \int_0^{\operatorname{tg} \pi/8} \frac{h^2}{(\rho^2 + h^2)^2} \rho d\rho \\ &= 2 \int_0^{\operatorname{tg} \pi/8} \frac{h^4}{(h^2 u^2 + h^2)^2} u du = 2 \int_0^{\operatorname{tg} \pi/8} \frac{1}{(1+u^2)^2} u du = \int_0^{\operatorname{tg}^2 \pi/8} \frac{1}{(1+v)^2} dv \\ &= \int_1^{1+\operatorname{tg}^2 \pi/8} \frac{1}{v^2} dv = -\frac{1}{v} \Big|_1^{1+\operatorname{tg}^2 \pi/8} = 1 - \frac{1}{1+\operatorname{tg}^2 \pi/8} = \frac{\operatorname{tg}^2 \pi/8}{1+\operatorname{tg}^2 \pi/8} \\ &= \cos^2 \pi/8 \operatorname{tg}^2 \pi/8 = \sin^2 \pi/8 \end{aligned}$$

Esercizio 3.

1. Una scena consiste dell'intercapedine fra il piano $z=0$ considerato come un pavimento infinito, che chiamiamo A_0 , ed un soffitto infinito dato dal piano $z=1$, che chiamiamo A_1 . Siano x_0, y_0 due numeri reali. Ovviamente il fattore di forma differenziale di A_0 (ossia il fattore di forma $F_{dA_1-A_0}$ del pavimento visto da un elemento infinitesimo di area dA_1 del soffitto) rispetto al punto $V=(x_0, y_0, 1)$ deve valere 1, perché il soffitto copre l'intero angolo solido nel semispazio frontale. Verificare questo fatto usando la formula integrale per il fattore di forma differenziale $F_{dA_1-A_0}$.
2. Ora la scena consiste in una stanza cilindrica il cui pavimento è il disco di raggio k centrato nell'origine nel piano $z=0$, che chiamiamo $A_0(k)$, ed il soffitto, che chiamiamo $A_1(k)$, giace nel piano $z=1$. Calcolare il fattore di forma differenziale da un punto del soffitto a $A_0(k)$, usando la formula integrale per il fattore di forma differenziale $F_{dA_1-A_0}$. Calcolare il limite di questo risultato per k che tende ad infinito e verificare che coincide con il caso limite della parte precedente (intercapedine infinita).

Svolgimento.

Parte 1. Il fattore di forma differenziale è

$$F = \int_{A_0} \frac{\cos \theta_0 \cos \theta_1}{\pi r^2} H_{01} d\sigma$$

dove il fattore di occlusione H_{01} vale sempre 1 perché non ci sono elementi che si frappongono, r è la lunghezza del segmento da $(x_0, y_0, 1)$ al punto generico $(x, y, 0)$ in A_0 , e gli angoli θ_0 e θ_1 sono quelli che questo segmento forma con il versore normale $(0,0,1)$ al

piano $z=0$ ed il versore normale $(0,0,-1)$ al piano $z=1$, rispettivamente. È quindi chiaro che $\cos \theta_0 = \cos \theta_1 = 1/r$, e quindi

$$F = \int_{\{z=0\}} \frac{1}{\pi r^4} dx dy = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{1}{\pi (1 + (x - x_0)^2 + (y - y_0)^2)^2} dx dy$$

Per l'invarianza per traslazione, l'ultimo integrale è costante rispetto a x_0, y_0 . Quindi

$$F = \int_0^{2\pi} \int_0^{\infty} \frac{1}{\pi (1 + \rho^2)^2} \rho d\rho d\theta$$

Ponendo $u = \rho^2$ si trova

$$F = \int_0^{\infty} \frac{1}{(1+u)^2} du = \int_1^{\infty} \frac{1}{u^2} du = 1$$

Parte 2. Procedendo come sopra, ora abbiamo

$$F = \int_{A_0(k)} \frac{1}{\pi r^4} dx dy = \int_0^{2\pi} \int_0^k \frac{1}{\pi (1 + \rho^2)^2} \rho d\rho d\theta$$

Ponendo $u = \rho^2$ si trova

$$F = \int_0^{k^2} \frac{1}{(1+u)^2} du = \int_1^{1+k^2} \frac{1}{u^2} du = \left| -\frac{1}{u} \right|_1^{1+k^2} = 1 - \frac{1}{1+k^2}$$

Chiaramente il limite per k che tende a infinito è 1.

Esercizio 4. Si consideri una ciotola K di forma emisferica di raggio 1 appoggiata nel suo punto più basso all'origine, ossia tangente all'origine al piano $\{z=0\}$ e contenuta nel semispazio superiore $\{z \geq 0\}$: quindi il bordo della ciotola è un cerchio J sul piano $\{z=1\}$. Poiché K è convessa, il fattore di forma dalla ciotola a sé stessa è positivo.

1. Si mostri che l'equazione che definisce K è $x^2 + y^2 + (z-1)^2 = 1$, con $z \leq 1$.
2. Si calcoli l'integrale del fattore di forma differenziale F_{0K} dall'origine a K e si mostri che vale $\frac{1}{2}$ (*suggerimento: si immagini di completare la scena in modo che sia uno spazio chiuso, ad esempio una sfera intera: se seguiamo questo suggerimento aggiungiamo una semisfera con bordo J , che chiamiamo C*).
3. Sia $v=(x_v, y_v, z_v)$ un altro punto di K . Si scriva esplicitamente l'integrale del fattore di forma differenziale da questo punto a K , in modo che l'integrando sia espresso come funzione solo delle coordinate x_v, y_v, z_v di v e x_p, y_p, z_p di un punto p che varia in K .

4. Sia p un punto del bordo J della ciotola. Utilizzando opportuni ragionamenti basati su simmetrie, si dimostri che il fattore di forma differenziale F_{pK} da questo punto alla ciotola è $\frac{1}{2}$.

Svolgimento.

Parte 1. Ovvio: la semisfera è la metà inferiore della sfera di raggio 1 e centro $(0,0,1)$.

Parte 2. Si consideri la restrizione di K al piano xz . I punti terminali di K in questo piano sono $(\pm 1,1)$. Sia θ l'angolo di deviazione dei punti di K dalla direzione verticale (misurata dall'origine): quindi per i punti di K la deviazione è $\pi/4 \leq \theta \leq \pi/2$. Completiamo K ad una sfera intera, e chiamiamo C la semisfera così aggiunta, il cui bordo è J . Chiaramente, per i punti di C , si ha $0 \leq \theta \leq \pi/4$. Per la conservazione dell'angolo solido totale, il fattore di forma differenziale è $F_{0K} = 1 - F_{0C}$ (la somma di tutti i fattori di forma differenziali da un punto a tutta la scena

è il 100% dell'angolo solido totale, ossia 1).

Il calcolo di F_{0C} è analogo a quello svolto negli esercizi precedenti: si può ricavare con il metodo di Nusselt oppure con il calcolo diretto dell'integrale di superficie,

$$\begin{aligned} F_{0C} &= \int_C \frac{\cos \theta}{\pi h^2} d\sigma = \int_0^{2\pi} \int_0^{\pi/4} \frac{\cos \theta}{\pi h^2} h^2 \sin \theta d\theta d\varphi = \int_0^{2\pi} \int_0^{\pi/4} \sin 2\theta d\theta d\varphi = \frac{1}{2} \int_0^{2\pi} \sin \theta d\theta \\ &= \frac{1}{2} (1 - \cos \pi/2) = \frac{1}{2} (1 - \cos^2 \pi/4 + \sin^2 \pi/4) = \sin^2 \pi/4 = \frac{1}{2} \end{aligned}$$

Pertanto, $F_{0K} = 1 - F_{0C} = 1 - \sin^2(\pi/4) = 1 - \frac{1}{2} = \frac{1}{2}$.

Parte 3. L'integrale è

$$F_{vK} = \int_K \frac{\cos \theta_v \cos \theta_p}{\pi r^2} H_{01} d\sigma(p)$$

dove il fattore di occlusione H_{01} vale costantemente 1, r è la distanza fra v e p , ossia la lunghezza del segmento s che li unisce, in altre parole

$$r^2 = (x_p - x_v)^2 + (y_p - y_v)^2 + (z_p - z_v)^2 \quad (1)$$

e θ_v e θ_p sono gli angoli che le normali a K in v e p (ciascuna coincidente con il rispettivo versore radiale, diretto verso il centro $(0,0,1)$) formano con il segmento s . Chiamiamo o il centro $(0,0,1)$ e consideriamo il triangolo $T_K = \{o, v, p\}$. I due lati che si incontrano in o hanno lunghezza 1 (pari al raggio!): ossia, il triangolo è isoscele. Pertanto ciascun coseno nell'integrando è uguale alla lunghezza della proiezione del corrispondente lato di lunghezza 1 sopra il terzo lato s che ha lunghezza r , ed evidentemente, visto che il triangolo è isoscele, le due proiezioni sono di lunghezza uguale. Quindi si ha

$$\cos \theta_v = \cos \theta_p = \frac{1}{2} \quad (2)$$

Ora, in base a (2), l'integrale diventa

$$F_{vK} = \frac{1}{4} \int_K \frac{1}{\pi r^2} d\sigma(p)$$

Per scrivere esplicitamente l'integrale superficiale ormai resta solo da rammentare quanto vale la misura $d\sigma$. In base alla parte 1, la superficie è parametrizzabile come

$$\sigma(x_p, y_p) = (x_p, y_p, 1 - \sqrt{1 - x_p^2 - y_p^2})$$

(davanti al radicale abbiamo scelto il segno meno perché $0 \leq z_p \leq 1$). Per questa parametrizzazione, se poniamo

$$f(x_p, y_p) = 1 - \sqrt{1 - x_p^2 - y_p^2}$$

il fattore di dilatazione dell'area (come già osservato nella Sottosezione [Fattore locale di ingrandimento delle aree nella parametrizzazione di una superficie, e calcolo dell'area di una superficie parametrica](#) della Sezione [Mappa di rilievo](#) del Capitolo [Rendering con mappe parametriche](#)) è

$$\left\| \frac{\partial \sigma}{\partial x_p} \times \frac{\partial \sigma}{\partial y_p} \right\| = \sqrt{1 + \|\nabla f\|^2} = \sqrt{1 + \frac{x_p^2 + y_p^2}{1 - x_p^2 - y_p^2}} = \sqrt{\frac{1}{1 - x_p^2 - y_p^2}}$$

Pertanto,

$$\begin{aligned} F_{vK} &= \frac{1}{4} \int_K \frac{1}{\pi r^2} \sqrt{\frac{1}{1 - x_p^2 - y_p^2}} dx_p dy_p \\ &= \frac{1}{4\pi} \int_K \frac{1}{(x_p - x_v)^2 + (y_p - y_v)^2 + (z_p - z_v)^2} \sqrt{\frac{1}{1 - x_p^2 - y_p^2}} dx_p dy_p \end{aligned}$$

dove abbiamo usato l'espressione di r^2 data da (1).

Variante per l'espressione dell'integrale. Supponiamo di chiudere la ciotola K con un coperchio dato da un disco B nel piano $\{z=1\}$. Allora, come prima, si ha $F_{vK} = 1 - F_{vB}$ perché gli angoli solidi che K e B sottendono quando visti dal punto v hanno per somma il 100% dell'angolo solido totale. Pertanto basta calcolare F_{vB} . In questo calcolo, l'integrale è sulla proiezione di B sul piano xy : infatti, le componenti del punto $q=(x_q, y_q, 1)$ che varia in B sono tutte e sole quelle che verificano $x_q^2 + y_q^2 \leq 1$. L'angolo θ_p è quello fra il segmento \mathbf{s} e la normale a B diretta verso il basso, ossia $(0,0, -1)$: pertanto $\cos \theta_p$ è la proiezione di \mathbf{s} sull'asse z , ossia

$$\cos \theta_p = 1 - z_v.$$

Invece l'angolo θ_v è, come prima, quello fra il segmento $\mathbf{s} = \mathbf{vq} = \mathbf{q} - \mathbf{v}$ e la normale interna a K al punto v , che è orientata come il raggio di K , verso il centro $(0,0,1)$. Si consideri

allora il triangolo $T_B = \{o, v, q\}$. Ora questo triangolo non è più isoscele, ma di nuovo, come prima, abbiamo

$$\cos \theta_v = \rho$$

dove ρ è la lunghezza del segmento da v al piede w della proiezione di o su \mathbf{s} . Calcoliamo questa lunghezza. Poiché il segmento da v a o ha lunghezza 1, è chiaro che la proiezione ortogonale è il prodotto scalare fra il vettore applicato $\mathbf{vo} = o-v$ ed il vettore applicato $\mathbf{vp} = p-v$. Ora, $o-v = (-x_v, -y_v, 1-z_v)$, e quindi il prodotto scalare vale

$$\rho = \langle o-v, p-v \rangle = -x_v(x_p - x_v) - y_v(y_p - y_v) + (1-z_v)(z_p - z_v) \quad (3)$$

Seconda variante per l'espressione dell'integrale. Possiamo immaginare di chiudere la ciotola K con un emisfero superiore C , in maniera da completare K ad una sfera. Il valore di $\cos \theta_p$ e di $\cos \theta_v$ si calcolano come nel caso in cui il punto finale q di $\mathbf{s} = \mathbf{vq}$ giaceva in K (il significato di entrambi gli angoli ritorna ad essere quello della deviazione angolare di \mathbf{s} rispetto ai versori normali interni alla sfera). Lasciamo i calcoli al lettore.

Parte 4. Basta completare la semisfera ad una sfera, aggiungendo una semisfera superiore. Un punto sul bordo J , per simmetria, ha lo stesso fattore di forma verso la semisfera bassa K e la semisfera alta.

SOTTOSTRUTTURAZIONE

Più fine è la suddivisione in elementi, migliore è il risultato ottenuto, ma il prezzo da pagare è un notevole incremento del tempo di calcolo perché il numero di fattori di forma è il quadrato del numero degli elementi, ed il calcolo di ciascun fattore di forma può essere laborioso. Per evitare questo incremento quadratico del numero dei fattori di forma, alcuni studiosi hanno proposto una suddivisione variabile degli elementi, basata sul ripartirli ulteriormente in sottoelementi laddove la radiosità varia più rapidamente. Per velocizzare il procedimento, i sottoelementi non vengono trattate come gli elementi originari. Infatti, per ciascun sottoelemento s si calcola solamente il fattore di forma F_{s-j} da s verso ciascuno degli elementi j originari (usando la tecnica del semicubo), mentre i fattori di forma dagli elementi ai sottoelementi non vengono calcolati. In tal modo miglioriamo il calcolo della distribuzione di energia ricevuta dagli elementi, considerando la geometria della ricezione di luce in modo più fine, spezzando le aree di ricezione laddove l'intensità è elevata, mentre non raffiniamo la distribuzione dell'emissione. Dopo che un elemento è stato suddiviso, i suoi fattori di forma verso ogni altro elemento vengono sostituiti con stime più accurate, e precisamente con le medie dei fattori di forma dei suoi sottoelementi, pesate in proporzione alle aree:

$$F_{ij} = \frac{1}{A_j} \sum_{s=1}^m F_{sj} A_s$$

La radiosità di ogni sottoelemento s di un elemento i viene calcolata riscrivendo il sistema lineare della radiosità (presentato nella sezione [L'equazione della radiosità](#)) rispetto ai sottoelementi, ma mantenendo per ciascun sottoelemento il coefficiente di riflessione dell'elemento da cui proviene (i coefficienti di riflessione per gli elementi originarie sono assegnati in fase di modellazione: non abbiamo a disposizione una stima più precisa per quelli dei sottoelementi). Di solito viene anche mantenuta per il sottoelemento la componente E_i di energia luminosa creata internamente che vale per l'elemento da cui esso proviene, perché si è interessati ad una migliore distribuzione dell'energia ricevuta e non di quella emessa (e quindi normalmente non si sottostrutturano le sorgenti di luce), ma se si decide di spezzare in sottoelementi anche le sorgenti di luce si può usare, per ciascun elemento s , un suo valore specifico E_s di energia creata.

Quindi abbiamo ora una stima più raffinata anche della radiosità dei sottoelementi, come media pesata in proporzione alle aree:

$$B_s = E_i + \rho_i \sum_{j=1}^n B_j F_{s-j}$$

Questa stima è stata ottenuta a partire dai valori già noti di B_j e quindi non richiede la risoluzione di un sistema lineare della radiosità a dimensione più elevata; peraltro, essa tiene conto in modo più accurato di come la illuminazione si distribuisce sui sottoelementi, ma non di come questi la irradiano a propria volta in modo più appropriato: per il calcolo della radiosità degli altri elementi si usa sempre B_i , non le B_s . L'algoritmo procede per iterazione, suddividendo gli elementi nelle zone dove la radiosità è maggiore, fino a quando il miglioramento ottenuto dalla suddivisione non raggiunge una soglia sufficientemente piccola prefissata. Gli elementi ottenuti da questo processo vengono usate per determinare la radiosità nei vertici.

Il fatto di determinare le radiosità ai singoli vertici è particolarmente vantaggioso per come è strutturato l'algoritmo di shading di [Gouraud](#), che si basa su metodi di interpolazione a partire dai dati sui vertici (rammentiamo che qui non si usa l'algoritmo di [Phong](#) perché la scena è puramente diffusiva, ovvero con coefficienti di riflessione pari a zero).

RICHIAMI SUL SISTEMA LINEARE DELLA RADIOSITÀ

Il sistema lineare della radiosità ricavato precedentemente (nella sezione sull'[Equazione della radiosità](#)) è del tipo $Mb = e$, dove e è il vettore dell'energia (o meglio potenza) luminosa creata dal singolo elemento (luce propria) e b è il vettore della radiosità, cioè dell'energia luminosa totale diffusa (quella creata o riflessa).

Se ρ_i indica il coefficiente di riflettività dell'elemento numero i e F_{ii} è il fattore di forma dall'elemento i a se stesso (che misura quale frazione dell'energia emessa dall'elemento i ritorna a quell'elemento), si ha:

$$0 \leq \rho_i < 1 \quad \text{e} \quad 0 \leq F_{ii} \leq 1$$

Abbiamo precedentemente visto che la matrice M del sistema lineare della radiosità verifica:

$$M_{ii} = 1 - \rho_i F_{ii} \quad i = 1, \dots, n$$

Abbiamo già osservato, nella sezione "[L'equazione della radiosità](#)", che $\rho_i < 1$ e $\rho_i F_{ii} < 1$. Quindi $0 < M_{ii} \leq 1$ per ogni i : i termini diagonali di M non si annullano. Perciò il sistema $Mb=e$, ossia

$$M_{ii}b_i + \sum_{j=1, \dots, n, j \neq i} M_{ij}b_j = e_i \quad i = 1, \dots, n$$

diventa :

$$b_i = - \sum_{j=1, \dots, n, j \neq i} (M_{ij}/M_{ii})b_j + e_i/M_{ii} \quad i = 1, \dots, n \quad (1)$$

Questo sistema si può risolvere con metodi iterativi.

SOLUZIONE CON METODI ITERATIVI: RILASSAMENTO DI JACOBI

Il più semplice metodo iterativo è quello di Jacobi: si parte con qualche vettore iniziale $b^{(0)}$ al posto della soluzione b esatta da trovare, e si sostituisce nel lato destro di (1), per ottenere una prima approssimazione iterativa:

$$b_i^{(1)} = - \sum_{j=1, \dots, n, j \neq i} (M_{ij}/M_{ii}) b_j^{(0)} + e_i/M_{ii} \quad i = 1, \dots, n \quad (2)$$

(rammentiamo che gli elementi diagonali M_{ii} sono non nulli).

Le componenti di $b^{(1)}$ vengono così determinate una dopo l'altra, in un primo ciclo (per i da 1 a N) che chiameremo la prima iterazione, ovvero il primo ciclo di iterazione. Poi si procede in modo analogo, ciclo dopo ciclo; alla m -sima iterazione determiniamo $b^{(m)}$ da

$$b_i^{(m)} = - \sum_{j=1, \dots, n, j \neq i} (M_{ij}/M_{ii}) b_j^{(m-1)} + e_i/M_{ii} \quad (3)$$

Si dice che il metodo è convergente se i $b^{(m)}$ convergono ad un vettore limite b , che in tal caso è necessariamente soluzione del sistema (1).

Si noti che tutte le componenti di $b^{(m)}$ vengono aggiornate insieme nello stesso ciclo, utilizzando tutte le componenti di $b^{(m-1)}$: quindi il metodo di Jacobi richiede di conservare in memoria simultaneamente due diversi aggiornamenti del vettore della radiosità: $b^{(m)}$ e $b^{(m-1)}$.

Metodi di questo tipo si chiamano anche *metodi di rilassamento*.

Ora ne vedremo due varianti più efficienti.

RILASSAMENTO DI GAUSS-SEIDEL APPLICATO AL SISTEMA LINEARE DELLA RADIOSITÀ

Il metodo iterativo di Gauss-Seidel è simile al metodo di Jacobi, tranne che, quando si svolge il ciclo di iterazione (3), le n equazioni vengono anche qui risolte una dopo l'altra in ordine progressivo (per i da 1 a n), ma, nella i -esima equazione, al membro di destra si utilizzano ove possibile le componenti già precedentemente aggiornate (nelle equazioni precedenti del ciclo, da 1 a $i-1$) del vettore $b^{(m)}$ invece che quelle di $b^{(m-1)}$. Per le componenti non ancora aggiornate, cioè da $i+1$ a n , si continua a utilizzare $b^{(m-1)}$. Quindi invece di (2) abbiamo:

$$b_i^{(m)} = -\sum_{j=1}^{i-1} \frac{M_{ij}}{M_{ii}} b_j^{(m)} - \sum_{j=i+1}^n \frac{M_{ij}}{M_{ii}} b_j^{(m-1)} + \frac{e_i}{M_{ii}} \quad (4)$$

In una prossima sezione dal titolo [Convergenza del metodo di Gauss-Seidel](#) mostreremo che il metodo di Gauss-Seidel è convergente se la matrice M ha la seguente proprietà di predominanza diagonale stretta:

Definizione. Una matrice $n \times n$ M è a predominanza diagonale stretta se, per ogni $i = 1, \dots, n$ si ha

$$|M_{ii}| > \sum_{j=1, \dots, n, j \neq i} |M_{ij}|$$

Lemma. La matrice M del sistema (1) della radiosità è a predominanza diagonale stretta.

Dimostrazione. Abbiamo visto nella sezione dal titolo [L'equazione della radiosità](#) che, per $1 \leq i \leq N$, si ha

$$M_{ii} = 1 - \rho_i F_{ii}.$$

Inoltre, per $1 \leq j \leq n$, si ha $\sum_{j=1, \dots, n} F_{ij} = 1$ (qui supponiamo che anche lo sfondo sia un elemento, e quindi che l'area totale coperta da tutti gli elementi vista dall'elemento numero i sia l'intero angolo solido del semispazio anteriore, di π steradiani).

Sempre nella stessa sezione abbiamo anche visto che, per $i \neq j$, si ha $M_{ij} = -\rho_i F_{ij}$. Perciò $M_{ii} = 1 - \rho_i F_{ii} = \sum_{j=1, \dots, n} F_{ij} - \rho_i F_{ii}$. Ma $\rho_i < 1$ e $M_{ii} > 0$ per ogni i . Quindi

$$|M_{ii}| > \sum_{j=1, \dots, n} \rho_i F_{ij} - \rho_i F_{ii} = \sum_{j=1, \dots, n, j \neq i} \rho_i F_{ij} = \sum_{j=1, \dots, n, j \neq i} |M_{ij}|,$$

il che completa la dimostrazione.

ANALISI DEL METODO DI GAUSS-SEIDEL

1. Convergenza di metodi di rilassamento

Dato un operatore lineare M su C^n , i metodi di rilassamento sono mirati a trasformare il problema lineare

$$Mb = e$$

in un problema equivalente di punto fisso che sia risolvibile in maniera rapida ed efficiente. Fissata una base in C^n , M si identifica con una matrice $n \times n$. La trasformazione del problema si ottiene attraverso uno spezzamento della matrice M :

$$M = N - P$$

con N invertibile. Così si ottiene:

$$Nb = Pb + e$$

$$b = N^{-1}Pb + N^{-1}e \quad (5)$$

Il problema $Mb=e$ è quindi trasformato nel problema di punto fisso

$$b = Qb + a \quad (5')$$

dove:

$$a = N^{-1}e \quad e \quad Q = N^{-1}P \quad (5'')$$

La matrice Q è chiamata la *matrice di iterazione*.

Ora, per ogni M c'è un unico spezzamento $M = N - P$ con N triangolare inferiore e P triangolare strettamente superiore (cioè con zeri sulla diagonale). Sappiamo già che gli elementi diagonali di N sono non nulli e quindi N è invertibile.

Il problema di punto fisso viene quindi affrontato con l'iterazione

$$b^{(k+1)} = Q b^{(k)} + a \quad (6)$$

ossia

$$b^{(k+1)} = N^{-1}P b^{(k)} + N^{-1}e \quad (6')$$

a partire da un vettore iniziale $b^{(0)}$ assegnato. Si noti che ora (5) diventa

$$Qb = b - N^{-1}e \quad (5''')$$

Affinché questo approccio sia vantaggioso occorre che il sistema lineare

$$N y = P b^{(k)} + e$$

(dove il termine di destra è un vettore noto e y è un vettore incognito) sia risolvibile per y in maniera rapida, cioè con un tempo di calcolo trascurabile rispetto a quello richiesto per la risoluzione del problema originale $Mb = e$. Per questo si cerca di ridursi al caso in cui l'operatore invertibile N sia espresso da una matrice triangolare (inferiore o superiore), perché in questo caso la soluzione avviene per sostituzioni successive di ogni equazione del sistema lineare nella precedente. Qui abbiamo scelto N triangolare inferiore.

Ricordiamo che il raggio spettrale di un operatore lineare T su C^n è definito da

$$\rho(T) = \limsup_{k \rightarrow \infty} \|T^k\|^{1/k}$$

E' ovvio che $\rho(T) \leq \|T\|$, perché $\|UV\| \leq \|U\| \|V\|$ per ogni coppia di operatori lineari su C^n .

Ora proveremo un teorema generale sulla convergenza di metodi di rilassamento: **se la matrice Q ha raggio spettrale minore di 1, allora l'errore tende a zero qualunque sia il vettore iniziale $b^{(0)}$** (e vedremo dopo che questa condizione è anche necessaria per la convergenza), *ma può accadere che, per specifiche scelte del vettore iniziale, l'errore tenda a zero senza che la matrice abbia raggio spettrale minore di 1*. Ciò accade quando definitivamente il nucleo delle matrici Q^k include $b^{(0)} - b$ (come si vede dalla identità (6'') più sotto; di solito si sceglie $b^{(0)} = 0$, quindi la condizione è che il punto fisso b sia nel nucleo delle matrici Q^k da un certo k in poi). **D'altra parte, se l'errore tende a zero per ogni vettore iniziale $b^{(0)}$, allora Q deve avere raggio spettrale minore di 1.**

Ecco la **dimostrazione**:

consideriamo l'incremento alla k -esima iterazione ($k > 0$),

$$\delta^{(k)} = b^{(k)} - b^{(k-1)}$$

Da (6') e (5''') si ottiene per questi incrementi la relazione di ricorrenza

$$\delta^{(k+1)} = b^{(k+1)} - b^{(k)} = Q b^{(k)} - Q b^{(k-1)} = Q \delta^{(k)}$$

e quindi

$$\delta^{(k+1)} = Q \delta^{(k)} = Q^2 \delta^{(k-1)} = \dots = Q^k \delta^{(1)} \quad (6''')$$

Dalla definizione di raggio spettrale,

$$\limsup_{k \rightarrow \infty} \|Q^k\|^{1/k} = \rho(Q)$$

segue che, se $\rho(Q) < 1$, si ha $\|\delta^{(k)}\| = \|Q^k \delta^{(1)}\| \leq \rho(Q)^k$, e quindi

$$\|\delta^{(k)}\| = \|Q^k \delta^{(1)}\| \leq C \rho(Q)^k$$

per qualche costante $C > 0$. Allora, poiché $\rho(Q) < 1$, la serie $\sum \delta^{(k)}$ converge.

D'altra parte, la somma parziale $\sum_{k=1}^n \delta^{(k)}$ è la somma telescopica

$$\sum_{k=1}^n \delta^{(k)} = \sum_{k=1}^n (b^{(k)} - b^{(k-1)}) = b^{(n)} - b^{(0)}$$

e quindi la convergenza della serie equivale all'esistenza del limite (finito)

$$b = \lim_{k \rightarrow \infty} b^{(k)}.$$

Si osservi che, definendo l'errore alla k-sima iterazione ($k > 0$) come

$$\eta^{(k)} = b^{(k)} - b$$

analogamente a prima e grazie anche a (5''') si ottiene per questi errori la relazione di ricorrenza

$$\eta^{(k+1)} = b^{(k+1)} - b = N^{-1}P b^{(k)} + N^{-1}e - b = Q b^{(k)} - Qb = Q\eta^{(k)}$$

E' da qui che segue cio' che abbiamo accennato poco sopra: se $\eta^{(0)} = b^{(0)} - b$ è nel nucleo di $\bigcap_{j=k+1}^{\infty} Q^j$ allora $0 = \eta^{(k+1)} = \eta^{(k+2)} = \eta^{(k+3)} \dots$

Viceversa, consideriamo il caso $\rho(Q) \geq 1$. Rammentiamo che il raggio spettrale di un operatore su uno spazio vettoriale a dimensione finita è il massimo dei moduli dei suoi autovalori. Sia v un autovettore di Q corrispondente all'autovalore di massimo modulo: questo modulo vale quindi

$\rho(Q) \geq 1$, e $Q^k v = \rho(Q)^k v$ non converge (anzi, se $\rho(Q) > 1$, la successione dei vettori $Q^k v$ diverge al crescere di k). Pertanto, se scegliamo $b^{(0)} = v$, vediamo che il procedimento iterativo non converge.

Abbiamo dimostrato:

Teorema 1. *Condizione necessaria e sufficiente affinché il metodo iterativo (6) converga per ogni scelta di vettore iniziale è che $\rho(Q) < 1$.*

2. Metodo di Gauss-Seidel (metodo delle sostituzioni successive)

La decomposizione adottata nel metodo di Gauss-Seidel è la seguente: $M = L + D + U$, dove D è una matrice diagonale (la diagonale di M) ed L e U sono la parte triangolare (strettamente) inferiore e superiore di M . Allora poniamo $N = L+D$ e $P=-U$, ed osserviamo che la matrice N è triangolare (inferiore), e quindi ha gli stessi autovalori di D , cioè ha per autovalori gli elementi diagonali M_{ii} di M . Ma abbiamo visto nella sezione dal titolo *Equazione della radiosità* che $M_{ii} > 0$ per ogni $i = 1, \dots, n$: quindi N è invertibile.

Ora, nel calcolo di $b_i^{(k+1)}$, utilizziamo al secondo membro i valori di $b_j^{(k+1)}$ invece di $b_j^{(k)}$ per $i < j$ (e per comodità di scrittura portiamoli al primo membro). Allora l'iterazione

$$Nb^{(k+1)} = Pb^{(k)} + e$$

assume la forma

$$\begin{array}{lcl} M_{11} b_1^{(k+1)} & = & -M_{12} b_2^{(k)} - M_{13} b_3^{(k)} \dots - M_{1n} b_n^{(k)} + e_1 \\ M_{21} b_1^{(k+1)} + M_{22} b_2^{(k+1)} & = & -M_{23} b_3^{(k)} \dots - M_{2n} b_n^{(k)} + e_2 \\ \dots & & \dots \\ \dots & & \dots \end{array} \quad (7)$$

$$M_{n1}b_1^{(k+1)} + M_{n2}b_2^{(k+1)} + \dots + M_{nn}b_n^{(k+1)} = e_n$$

Come si può osservare, in questo metodo (della sostituzione successive) il valore aggiornato di ogni singola componente $b_j^{(k+1)}$ ($j=1, \dots, n$) viene utilizzato immediatamente per il calcolo delle componenti successive relative alla stessa iterazione, mentre in generale, nella risoluzione del sistema (6) con il metodo di Jacobi (introdotto nella sezione dal titolo *Metodi iterativi*), i valori aggiornati di tutte le componenti vengono sostituiti simultaneamente alla fine dell'iterazione.

3. Convergenza del metodo di Gauss-Seidel

Consideriamo un sistema lineare $Ax=b$, e le sue iterazioni con le sostituzioni successive nel senso di (7) (dove invece di una matrice quadrata generica A abbiamo considerato, come in (4), la matrice M della radiosità). Se il procedimento iterativo converge, la soluzione $x = (x_1, \dots, x_n)$ del sistema (4) (qui scritto come in (7)) verifica quindi:

$$x_i = -\sum_{j=1}^{i-1} \frac{a_{ij}}{a_{ii}} x_j - \sum_{j=i+1}^n \frac{a_{ij}}{a_{ii}} x_j + \frac{b_i}{a_{ii}} \quad (i=1, \dots, n)$$

mentre l'approssimazione di ordine $k+1$ verifica

$$x_i^{(k+1)} = -\sum_{j=1}^{i-1} \frac{a_{ij}}{a_{ii}} x_j^{(k+1)} - \sum_{j=i+1}^n \frac{a_{ij}}{a_{ii}} x_j^{(k)} + \frac{b_i}{a_{ii}}.$$

Quindi gli errori

$$\eta_i^{(k)} = x_i^{(k)} - x_i$$

risolvono ad ogni iterazione il seguente sistema lineare:

$$\eta_i^{(k+1)} = -\sum_{j=1}^{i-1} \frac{a_{ij}}{a_{ii}} \eta_j^{(k+1)} - \sum_{j=i+1}^n \frac{a_{ij}}{a_{ii}} \eta_j^{(k)}$$

Perciò:

$$|\eta_i^{(k+1)}| \leq \sum_{j=1}^{i-1} \left| \frac{a_{ij}}{a_{ii}} \right| |\eta_j^{(k+1)}| + \sum_{j=i+1}^n \left| \frac{a_{ij}}{a_{ii}} \right| |\eta_j^{(k)}| \quad (i=1, \dots, n) \quad (7')$$

e quindi:

$$|\eta_i^{(k+1)}| \leq \sum_{j=1}^{i-1} \left| \frac{a_{ij}}{a_{ii}} \right| \|\eta^{(k+1)}\| + \sum_{j=i+1}^n \left| \frac{a_{ij}}{a_{ii}} \right| \|\eta^{(k)}\| \quad (i=1, \dots, n) \quad (7'')$$

Detto $M=M_k$ l'indice i per il quale il secondo termine della disuguaglianza è massimo, e scrivendo per brevità:

$$R_M = \sum_{j=1}^{M-1} \left| \frac{a_{Mj}}{a_{MM}} \right| \quad \text{e} \quad S_M = \sum_{j=M+1}^n \left| \frac{a_{Mj}}{a_{MM}} \right|$$

si ottiene una disuguaglianza uniforme rispetto all'indice i , ossia tale che il secondo membro non dipende più da i :

$$|\eta_i^{(k+1)}| \leq R_M \|\eta^{(k+1)}\| + S_M \|\eta^{(k)}\| \quad (i=1, \dots, n)$$

Quindi, scegliendo come norma la norma ℓ^∞ , ossia $\|\eta^{(k+1)}\| = \max_{i=1 \dots n} |\eta_i^{(k+1)}|$ otteniamo

$$\|\eta^{(k+1)}\| \leq R_M \|\eta^{(k+1)}\| + S_M \|\eta^{(k)}\| \quad (7'')$$

(rammentiamo che in dimensione finita tutte le norme sono equivalenti, si veda l'Appendice degli [Appunti di Algebra Lineare](#); ma anche senza questa scelta di norma, da (7') si deduce (7'') e quindi (7''') in qualsiasi scelta di norma, ad esempio nella consueta norma Euclidea). Pertanto

$$(1 - R_M) \|\eta^{(k+1)}\| \leq S_M \|\eta^{(k)}\|.$$

Se supponiamo A a predominanza diagonale stretta, cioè se

$$|a_{ii}| > \sum_{i \neq j} |a_{ij}|$$

si ha:

$$R_M + S_M < 1$$

Ne segue che

$$R_M < 1 \quad \text{e} \quad S_M < 1$$

e

$$S_M / (1 - R_M) < 1$$

In questa disuguaglianza, l'indice $M = M_k$ dipende dalla scelta di k . Però M varia solo nell'insieme finito da 1 a n , e quindi la costante $S_M / (1 - R_M)$ varia in un insieme finito di numeri tutti minori di 1, in base all'ultima disuguaglianza. Con abuso di notazione, indichiamo il suo valore massimo ancora con $S_M / (1 - R_M) < 1$. Allora

$$\|\eta_i^{(k+1)}\| \leq \frac{S_M}{1 - R_M} \|\eta^{(k)}\| \leq \left(\frac{S_M}{1 - R_M} \right)^{k+1} \|\eta^{(0)}\| \quad (8)$$

Ora, grazie alla disuguaglianza $S_M / (1 - R_M) < 1$, segue da (8) che $\lim_{k \rightarrow \infty} \eta_i^{(k+1)} = 0$ (ed osserviamo anche che da questo e dalla dimostrazione del precedente Teorema 1 segue che il raggio spettrale della matrice Q è minore di 1).

Abbiamo dimostrato dunque il seguente teorema.

Teorema 2. *Se la matrice A è a predominanza diagonale stretta allora il metodo di Gauss-Seidel è convergente.*

Poiché abbiamo dimostrato in un Lemma precedente che la matrice della radiosità è a predominanza diagonale stretta, questo dimostra il fatto seguente:

Corollario. *Il metodo di Gauss-Seidel, applicato al problema della radiosità, è convergente.*

RILASSAMENTO DI SOUTHWELL APPLICATO AL SISTEMA LINEARE DELLA RADIOSITÀ

1. Rilassamento di Southwell

Abbiamo visto nella sezione precedente che in ogni ciclo di iterazione, da ognuna delle n equazioni nell'ordine $j = 1, \dots, n$, il metodo di Gauss-Seidel calcola la componente j di una approssimazione del valore di radiosità di ogni singolo elemento.

Ora presentiamo il metodo di Southwell, che procede in ordine diverso: anch'esso calcola le varie componenti della nuova approssimazione risolvendo una equazione lineare alla volta, ma la scelta di j (cioè di quale equazione usare) è fatta in modo da rendere massima l'efficienza.

Quando si studiano scambi di radiosità questi paragoni sono più significativi se compiuti su valori di energia (emessi da un singolo elemento) piuttosto che su valori di radiosità (energia per unità di area).

Riscriviamo l'equazione della radiosità in termini di valore di energia per elemento e definiamo nuovi vettori β e ε come

$\beta_i = A_i b_i$ energia totale che lascia l'elemento i (emessa internamente o riflessa) (8')

$\varepsilon_i = A_i e_i$ energia emessa internamente dall'elemento i

Quindi riscriviamo l'equazione della radiosità, trasformando la matrice M del sistema lineare (introdotta nella sezione dal titolo *L'equazione della radiosità*) in una nuova matrice K , che rappresenta il sistema lineare dell'energia:

$$K\beta = \varepsilon, \quad \text{con} \quad K_{ij} = (A_i/A_j) M_{ij} \quad (9)$$

Grazie alla relazione di reversibilità $A_i F_{ij} = A_j F_{ji}$ (anch'essa introdotta nella sezione dal titolo *L'equazione della radiosità*), nel calcolo degli elementi della matrice K si possono far apparire i termini F_{ji} invece che gli F_{ij} che abbiamo usato nella matrice M . Quindi si ha:

$$K_{ij} = (A_i/A_j) M_{ij} = (A_i/A_j) (\delta_{ij} - \rho_i F_{ij})$$

Si osservi ora che, grazie alla relazione di reversibilità, per $i \neq j$, si ha

$$K_{ij} = -\rho_i F_{ji} \quad (10)$$

mentre, nel caso $i=j$, si ha $K_{ii} = 1 - \rho_i F_{i-i}$. Pertanto, per ogni i, j , si ottiene

$$K_{ij} = \delta_{ij} - \rho_i F_{ji} \quad (11)$$

ed inoltre

$$\text{se } F_{i-i} = 0, \text{ allora } K_{ii} = 1 \quad (12)$$

Il sistema lineare (0) della radiosità quindi diventa il seguente sistema lineare, che chiamiamo il sistema della potenza:

$$\beta_i = \varepsilon_i + \rho_i \sum_{j=1, \dots, n} \beta_j F_{ji} \quad (0 \text{ bis})$$

ossia il sistema lineare trasposto di (0): quello definito dalla matrice trasposta.

L'algoritmo di Southwell funziona calcolando una serie di approssimazioni ($\beta^{(k)}$) della soluzione β . Il vettore degli errori, o *vettore dei resti*, o anche *resto*, che ci permette di stimare la qualità dell'approssimazione, è:

$$r^{(k)} = \varepsilon - K\beta^{(k)} \quad (13)$$

Il fatto che $\beta^{(k)}$ converga alla soluzione esatta β equivale al fatto che il vettore resto $r^{(k)}$ tenda a 0.

Notiamo che

$$K\beta_i^{(k)} = \sum_j K_{ij} \beta_j^{(k)} = \sum_j (A_i/A_j) M_{ij} \beta_j^{(k)} = A_i \sum_j M_{ij} b_j^{(k)} = A_i (Mb^{(k)})_i$$

Allora dalla definizione di ε e β e da (9) segue che la i -sima componente del vettore resto alla k -sima iterazione è data da

$$r_i^{(k)} = A_i ((e_i - (Mb^{(k)})_i)) \quad (13')$$

L'algoritmo calcola ad ogni passo una nuova approssimazione $\beta^{(k+1)}$ ottenuta modificando un singolo componente del vettore $\beta^{(k)}$ in modo tale che una delle componenti di $r^{(k)}$ diventi nulla: resto 0 per quella componente. La componente risolta ad una data iterazione si dice *rilassata*. A tal fine, vediamo cosa comporta il fatto che ad ogni iterazione l'algoritmo si limiti a risolvere (ovvero a rilassare) solo una delle n equazioni del sistema. Supponiamo che alla iterazione $k+1$ venga risolta la equazione numero i : questo vuol dire, per (6'), che si trova un approssimante $b^{(k)}$ tale che

$$b_i^{(k+1)} = (N^{-1}P b^{(k)})_i + (N^{-1}e)_i.$$

Invece per le altre componenti $b_j^{(k+1)}$ coincide con $b_j^{(k)}$. Nel metodo di Southwell, l'equazione da rilassare ad ogni iterazione viene scelta come quella che alla iterazione precedente dava luogo alla componente del vettore resto maggiore delle altre (ossia, diciamo così, viene rilassata la componente con l'errore più grande).

Quindi la differenza dal metodo di Gauss-Seidel è che questo aggiorna in maniera progressiva tutte le componenti del vettore della soluzione approssimata alla iterazione precedente, mentre il metodo di Southwell sceglie nel vettore resto $r^{(k)}$ la coordinata i alla quale corrisponde la componente più grande e rilassa la corrispondente componente $\beta_i^{(k)}$ del vettore in modo da annullare $r_i^{(k)}$.

Esaminiamo l'andamento dell'errore di approssimazione nel metodo di Southwell. Supponiamo che, al passo k , la componente più grande del vettore resto sia $r_i^{(k)}$ e che al passo successivo diventi zero:

$$0 = r_i^{(k+1)} = \varepsilon_i - \sum_{j=1, \dots, n} K_{ij} \beta_j^{(k+1)} \quad (14)$$

Poiché modifichiamo soltanto la componente i si ha

$$\beta_j^{(k+1)} = \beta_j^{(k)} \quad \text{per } j \text{ diverso da } i \quad (15)$$

mentre per la componente i segue da (14) e da (15) che

$$\beta_i^{(k+1)} = \frac{1}{K_{ii}} \left(\varepsilon_i - \sum_{j \neq i} K_{ij} \beta_j^{(k)} \right) \quad (16)$$

Ossia, sempre da (14),

$$\beta_i^{(k+1)} = \frac{1}{K_{ii}} \left(\varepsilon_i - \sum_{j=1}^n K_{ij} \beta_j^{(k)} + K_{ii} \beta_i^{(k)} \right) = \beta_i^{(k)} + \frac{r_i^{(k)}}{K_{ii}} \quad (17)$$

Da (13) segue che

$$r^{(k+1)} - r^{(k)} = -K (\beta^{(k+1)} - \beta^{(k)}) \quad (18)$$

Da (18), grazie a (15) e (17), nella moltiplicazione riga per colonna all'ultimo membro entra in gioco solo la colonna i -sima della matrice K (perché solo la riga i -sima del vettore $\beta^{(k+1)} - \beta^{(k)}$ è non nulla) e si ottiene

$$r_j^{(k+1)} = r_j^{(k)} - (K_{ji} / K_{ii}) r_i^{(k)} \quad (19)$$

Ripetiamo quanto sopra in un breve sommario. L'algoritmo di Southwell calcola per iterazione una soluzione approssimata $\beta^{(k)}$ e stima l'errore di approssimazione, cioè il vettore resto $r^{(k)}$.

A questo punto l'algoritmo trova quale è la componente del vettore resto più grande in valore assoluto ed aggiorna quella componente in maniera che il resto corrispondente diventi nullo. In questo modo cambiano anche le altre componenti; l'algoritmo ripete l'iterazione su quella di grandezza maggiore e così via.

Durante ogni iterazione viene usata una sola colonna della matrice K , quindi dobbiamo calcolare una sola riga della matrice dei fattori di forma (grazie a (10), la matrice K , al di fuori della diagonale, è la trasposta della matrice dei fattori di forma). La scelta del vettore $\beta^{(0)}$ è arbitraria: infatti ora dimostriamo che l'algoritmo converge per qualunque scelta iniziale.

2. Convergenza del metodo di Southwell

Consideriamo la norma del vettore resto. Dobbiamo dimostrare che essa converge a zero quando iteriamo il procedimento all'infinito. Al posto della norma Euclidea

$$|||r^{(k)}||| = \left[\sum_{i=1, \dots, n} (r_i^{(k)})^2 \right]^{1/2}$$

è equivalente usare la norma ℓ^1 data da

$$||r^{(k)}|| = \sum_{i=1, \dots, n} |r_i^{(k)}|$$

perché

$$|||r^{(k)}||| \leq ||r^{(k)}|| \leq n^{1/2} |||r^{(k)}|||$$

grazie alla disuguaglianza di Cauchy-Schwarz (si vedano gli [Appunti di Algebra Lineare](#) (Appendice) o di [Analisi di Fourier](#)). (Rammentiamo comunque di nuovo che in dimensione finita tutte le norme sono equivalenti: si riveda l'Appendice delle [Appunti di Algebra Lineare](#).) Quindi vogliamo provare che

$$\lim_{k \rightarrow \infty} \|r^{(k)}\| = \lim_{k \rightarrow \infty} \sum_{i=1, \dots, n} |r_i^{(k)}| = 0$$

Da (19) abbiamo :

$$\|r^{(k+1)}\| = \sum_{j=1, \dots, n} \left| r_j^{(k)} - \frac{K_{ji}}{K_{ii}} r_i^{(k)} \right| \quad (20)$$

Poiché stiamo supponendo che alla iterazione k-sima la componente più grande del vettore $r^{(k)}$ sia la i-esima, cioè che sia questa la componente che viene rilassata alla iterazione k+1, segue che $r_i^{(k+1)} = 0$, in base a (13) o anche a (19). Quindi in realtà la somma in (20) è limitata solo a tutti gli indici j diversi da i.

Ora dalla disuguaglianza triangolare abbiamo

$$\|r^{(k+1)}\| \leq \sum_{j \neq i} |r_j^{(k)}| + \sum_{j \neq i} \left| \frac{K_{ji}}{K_{ii}} r_i^{(k)} \right| = \|r^{(k)}\| - |r_i^{(k)}| + |r_i^{(k)}| \sum_{j \neq i} \left| \frac{K_{ji}}{K_{ii}} \right| \quad (21)$$

Poiché la matrice K è a predominanza diagonale stretta, per ogni i, $1 \leq i \leq n$, si ha:

$$|K_{ii}| > \sum_{j \neq i} K_{ji} \quad (22)$$

Per $i=1, \dots, n$, poniamo

$$t = \max_{j \neq i} \sum \left| \frac{K_{ji}}{K_{ii}} \right|$$

In altre parole, $0 < t < 1$ per (22), e

$$\sum_{j=1, \dots, n, j \neq i} \left| \frac{K_{ji}}{K_{ii}} \right| < t \quad \text{per } 1 \leq i \leq n \quad (23)$$

Pertanto segue da (21) che:

$$\|r^{(k+1)}\| \leq \|r^{(k)}\| - (1-t) |r_i^{(k)}| \quad (24)$$

Per l'ipotesi del metodo di Southwell, all'iterazione k-sima la componente più grande del vettore r è la i-esima. Perciò:

$$\|r^{(k)}\| \leq n |r_i^{(k)}| \quad (25)$$

Quindi
$$(1-t) \left| r_i^{(k)} \right| \geq \frac{1-t}{n} \left\| r^{(k)} \right\| \quad (26)$$

e, sostituendo (26) in (24), otteniamo:

$$\left\| r^{(k+1)} \right\| \leq \left(1 - \frac{1-t}{n} \right) \left\| r^{(k)} \right\| \quad (27)$$

Osserviamo che, se si pone

$$T = 1 - \frac{1-t}{n}$$

applicando iterativamente (27) si ottiene

$$\left\| r^{(k+1)} \right\| \leq T^{k+1} \left\| r^{(0)} \right\| \quad (28)$$

Ma $T < 1$ poiché $0 < t < 1$, e quindi $\lim_{k \rightarrow \infty} T^{k+1} = 0$. Ne segue che $\lim_{k \rightarrow \infty} \|r^{(k)}\| = 0$, e perciò l'algoritmo converge per qualsiasi scelta delle condizioni iniziali.

3. Interpretazione fisica del procedimento di Southwell: luce emessa nell'ambiente

E' interessante interpretare i vari passi del metodo di Southwell quando esso viene applicato al problema della radiosità e vedere il significato fisico del rilassamento.

Per comodità, iniziamo il processo iterativo di Southwell assegnando al vettore approssimante valore zero (cioè tutte le componenti uguali a zero). In questo caso, in base a (13), il vettore iniziale dei resti è uguale al vettore dell'energia creata dall'elemento. La prima iterazione seleziona quindi la componente che corrisponde all'elemento che emette più luce propria. Questo primo passaggio porta ad una nuova stima dell'energia dell'elemento, calcolata in modo che il valore del resto diventi nullo per quell'elemento.

Assumendo per semplicità che tutti i fattori di forma da un elemento su sé stesso siano nulli, se indichiamo con i l'indice dell'elemento che viene rilassata durante l'iterazione $k+1$ (cioè quella con il resto più grande dopo l'iterazione k), la variazione del resto per ciascun elemento può essere espressa grazie a (19), (10) e (9) come

$$r_j^{(k+1)} = r_j^{(k)} - (K_{ji} / K_{ii}) r_i^{(k)} = r_j^{(k)} + \rho_j F_{ij} r_i^{(k)}$$

dove $j=1, \dots, n$, e i è l'indice dell'elemento di massima energia residua $r_i^{(k)}$ dopo l'iterazione k .

In altre parole, all'iterazione $k+1$, il resto $r_i^{(k)}$ dopo l'iterazione k viene distribuito fra tutti gli altri elementi in funzione dell'appropriato fattore di forma e del coefficiente di riflettività. Poiché questi resti sono differenze di energia, ogni iterazione modella il trasferimento di energia da un elemento all'ambiente, dovuto ad un'ulteriore riflessione della luce dai vari elementi: all'iterazione $k+1$ l'elemento i di massimo resto, ossia di massima energia immagazzinata, si libera della propria energia distribuendola nell'ambiente in proporzione ai fattori di forma. Come conseguenza di queste iterazioni di irraggiamento di energia, ciascuna degli altri elementi riceve e accumula energia luminosa, fino a che non viene il suo turno di emetterla: il suo turno arriva quando l'energia accumulata diventa così elevata da far sì che quel particolare elemento sia quello fra tutti con la massima quantità di energia luminosa immagazzinata.

VERSO UN RAFFINAMENTO PROGRESSIVO

La mole di calcoli dell'algoritmo di radiosità, se eseguito così come presentato fino a questo momento, è molto elevata. Sembra naturale domandarsi se è possibile procedere per approssimazioni successive, creando una prima immagine non accurata e migliorandola in seguito mediante un algoritmo incrementale. Con lo schema di calcolo che abbiamo stabilito per la radiosità questo procedimento non è realizzabile nel metodo di Southwell. Infatti i processi di iterazione richiedono che la nuova luminosità dell'elemento sia calcolata sommando tutte le energie entranti provenienti dagli altri elementi: non basta tener conto dell'emissione da parte di un solo elemento. Quindi, progredire nella stima della radiosità, deve essere eseguito l'intero ciclo di iterazione. Se ci sono n elementi, ogni ciclo di iterazione richiede la moltiplicazione di una matrice $n \times n$ con un vettore colonna di dimensione n , un procedimento che richiede n^2 moltiplicazioni, e quindi un tempo di esecuzione dell'ordine di $O(n^2)$; inoltre i fattori di forma tra tutte gli elementi devono essere precalcolati all'inizio ed essere memorizzati prima dell'esecuzione (perché il sistema (0 bis) richiede una somma su j che coinvolge la matrice trasposta F_{ji} , la quale quindi va calcolata, per ogni riga i , al variare di j , ossia usando per ogni elemento j un corrispondente z -buffer e non lo stesso per tutta la riga. Questo richiede un tempo e uno spazio di memoria dell'ordine di $O(n^2)$, perché ci sono n elementi. Comunque, quest'ultimo ostacolo si può aggirare grazie alla equazione di reversibilità.

In [M.F.Cohen, S.E.Chen, J.R.Wallace e D.P.Greenberg, *A Progressive Refinement Approach to Fast Radiosity Image Generation*, SIGGRAPH '89, in *Computer Graphics*, 23(3), luglio 1989] viene sviluppato un algoritmo per il raffinamento progressivo del calcolo della radiosità, senza la necessità di precalcolare e memorizzare tutti i fattori di forma: il sistema lineare viene risolto in modo progressivo, una riga per volta. Nel metodo iterativo fin qui presentato, il calcolo della radiosità dell'elemento i -esimo richiede il calcolo dei contributi di energia ricevuti dagli altri elementi. Ogni termine della i -esima equazione del sistema della radiosità esprime l'effetto sulla radiosità dell'elemento i dovuta all'emissione dell'elemento j per tutti i valori di j da 1 a n ; l'energia (o meglio potenza) contribuita all'elemento i dalla radiosità b_j dell'elemento j e poi riflessa nell'ambiente è

$$\rho_i b_j F_{ji} \quad (28')$$

Riassumendo, i metodi iterativi calcolano l'incremento di radiosità di un elemento grazie all'energia luminosa che esso raccoglie dagli altri. Abbiamo già osservato che tale procedimento è troppo oneroso.

EMISSIONE DI ENERGIA NELL'AMBIENTE INVECE DI ASSORBIMENTO

La revisione del procedimento, che porta ad un affinamento progressivo molto meno oneroso dell'immagine generata iterazione dopo iterazione, consiste nel reinterpretare il sistema lineare della potenza (0 bis) considerando la potenza emanata da un elemento verso il resto dell'ambiente, piuttosto che quella ricevuta: ed in effetti, scambiando i ruoli di i e j rispetto a prima (sezione *Equazione della radiosità*), ritroviamo proprio che il contributo alla radiosità b_j dell'elemento j dovuto all'emissione di radiosità dell'elemento i e poi riflesso nell'ambiente è $\rho_j b_i F_{ji}$, come in (28'), e questo conduce al sistema lineare trasposto di quello della radiosità.

Sfortunatamente questo procedimento richiede una maggiore mole di calcoli e di occupazione di memoria di quello originale (0) della radiosità: fissato l'elemento i , esso richiede la conoscenza del fattore di forma F_{ji} dall'elemento j all'elemento i , ed al variare di j ciascuno degli F_{ji} richiede un calcolo diverso, fatto su un diverso semicubo, quindi non con lo stesso z -buffer. Però qui possiamo utilizzare la relazione di reversibilità $A_i F_{ij} = A_j F_{ji}$ fra i fattori di forma, e quindi scrivere il contributo alla radiosità b_j dovuto a b_i (cioè la frazione di energia emessa dall'elemento i verso l'elemento j e da questo riflessa indietro verso il resto dell'ambiente) come

$$\rho_j b_i F_{ij} A_i/A_j$$

E' facile verificare che, con questa rinormalizzazione, dal sistema della potenza (0 bis) si ritorna al sistema della radiosità (0), ma adesso ne abbiamo reinterpretato il significato fisico: nel corso del k -simo ciclo di iterazione, la i -sima riga del sistema lineare della radiosità rappresenta l'emissione di energia lanciata nell'ambiente nel corso di quella iterazione dall'elemento i verso ciascuno degli altri elementi (incluso sé stesso se F_{i-i} è non nullo).

VANTAGGI NUMERICI DELL'APPROCCIO A EMISSIONE DI ENERGIA NELL'AMBIENTE

Il calcolo degli addendi a secondo membro di quest'ultima equazione al variare di j richiede solo la conoscenza dei fattori di forma F_{ij} calcolati usando un unico semicubo centrato al centro dell'elemento i -esimo. Se i fattori di forma dall'elemento i possono essere calcolati velocemente (ad esempio utilizzando uno z-buffer hardware), allora il tempo necessario per questo calcolo è breve; inoltre lo spazio di memoria occupata da F_{ij} può essere rilasciata non appena la radiosità emanata dall'elemento i verso l'elemento j viene lanciata nell'ambiente (cioè appena calcolato l'addendo j -simo). In questo modo, ad ogni passo dobbiamo avere a disposizione memoria solo per un singolo semicubo e per i suoi fattori di forma. L'algoritmo, la cui convergenza è stata dimostrata nel Teorema 2 (dimostrato nella sottosezione *Convergenza del metodo di Gauss-Seidel*), viene iterato fino a quando non si raggiunge una tolleranza prefissata.

Inizialmente la maggior parte dell'energia luminosa è concentrata in pochi elementi (le sorgenti luminose). Se il vettore di partenza $b^{(0)}$ dell'iterazione si pone uguale a 0, al primo passo avremo il primo iterato $b^{(1)}$ uguale al vettore e dell'energia creata dalle sorgenti: ossia vediamo solo le sorgenti ed il resto è buio. Al passo successivo dobbiamo cominciare ad aggiungere gli addendi della somma, che rappresentano, per ciascun elemento i , l'energia addizionale emessa dall'elemento i verso ciascuno degli altri e da questi riflessa indietro: in questa passo di iterazione, quindi, cominciano ad illuminarsi, oltre alle sorgenti di luce, anche gli elementi direttamente illuminati da esse. È chiaro che la convergenza diventa un pò più rapida se sviluppiamo il calcolo dapprima per la sorgente che ha più energia da emettere, e così via: questo significa che al primo passo ordiniamo gli elementi che sono sorgenti di luce in base alla luminosità (le altre, che non emettono luce propria, al primo passo non intervengono). Poi, via via, ad ogni ciclo di iterazione successivo scegliamo analogamente gli elementi da rilassare ordinandoli in base alla quantità di energia immagazzinata perché ricevuta dagli elementi che hanno precedentemente irradiato luce (in ciascuna fase di iterazione si rendono visibili gli elementi che ricevono luce direttamente dagli oggetti che erano illuminati alla fase precedente). In altre parole, non procediamo secondo l'ordine progressivo degli elementi, per i da 1 a n , bensì scegliamo volta per volta l'elemento di massima energia. Questo significa utilizzare, invece del metodo di Gauss-Seidel, il metodo di Southwell che è più efficiente.

STIMA DELLA RADIOSITÀ RESIDUA DA EMETTERE

Nella sezione *Rilassamento di Southwell* abbiamo studiato il vettore β dell'energia irradiata e le sue iterazioni $\beta^{(k)}$ nella soluzione iterativa del sistema lineare (9) dello scambio di energia fra gli elementi. Supponiamo che l'elemento con maggiore energia dopo la k-sima iterazione (cioè quella che nel corso dell'iterazione k+1 rilascia la sua energia accumulata) sia l'elemento i-esimo, e per semplicità facciamo l'ipotesi che non rifletta luce su sé stesso, cioè che il fattore di forma F_{i-i} sia zero. Allora, per (12), il corrispondente termine diagonale della matrice K del sistema lineare dello scambio di energia vale $K_{ii} = M_{ii} = 1$, e quindi, nel ciclo di iterazione k+1, l'incremento del vettore β dell'energia totale irradiata è dato dall'equazione (17):

$$\beta_i^{(k+1)} = \beta_i^{(k)} + r_i^{(k)} \quad (17 \text{ bis})$$

e, per j diverso da i, dall'equazione (15):

$$\beta_j^{(k+1)} = \beta_j^{(k)}.$$

Corrispondentemente si ha $r_i^{(k+1)} = 0$ (grazie a (14): rilassamento dell'elemento i) e, per j diverso da i,

$$r_j^{(k+1)} = r_j^{(k)} - K_{ji} r_i^{(k)} = r_j^{(k)} + \rho_j F_{ij} r_i^{(k)} \quad (29)$$

grazie all'equazione (19).

L'energia totale irradiata dall'elemento j al passo di iterazione k+1 è quindi la stessa che al passo precedente ($\beta_j^{(k+1)} = \beta_j^{(k)}$), ma c'è una quantità $r_j^{(k+1)} - r_j^{(k)}$ di energia addizionale pervenuta all'elemento in conseguenza del rilascio, in quella iterazione, dell'energia immagazzinata nell'elemento i fino alla iterazione precedente. Perciò è opportuno introdurre un nuovo vettore β' dell'energia totale già irradiata da ciascun elemento o in attesa di esserlo al successivo rilassamento:

$$\beta_j'^{(k)} = \beta_j^{(k)} + r_j^{(k)}, \quad j = 1, \dots, n \quad (30).$$

Questo vettore, un passo di iterazione dopo l'altro, dà una stima progressivamente più precisa dell'energia totale che verrà irradiata dagli elementi al termine del processo iterativo.

Segue ora da (30), (15) e (29) che, per ogni j diverso da i,

$$\beta_j'^{(k+1)} = \beta_j^{(k)} + r_j^{(k+1)} = \beta_j'^{(k)} + r_j^{(k+1)} - r_j^{(k)} = \beta_j'^{(k)} + \rho_j F_{ij} r_i^{(k)} \quad (31).$$

Ora ritorniamo alle radiosità, che sono le energie irradiate divise per le aree dei rispettivi elementi. Vogliamo una stima iterativa $b_i'^{(k)}$ della radiosità finale b_i , per ogni

elemento i. Essa è la somma della radiosità effettivamente rilasciata da quell'elemento fino al ciclo di iterazione k, che in base ad (8') è data da

$$b_i^{(k)} = \beta_i^{(k)} / A_i$$

e della radiosità immagazzinata ed ancora in attesa di essere rilasciata,

$$\Delta b_i^{(k)} = r_i^{(k)} / A_i .$$

Pertanto,

$$b_i'^{(k)} = b_i^{(k)} + \Delta b_i^{(k)} = (\beta_i^{(k)} + r_i^{(k)}) / A_i = \beta_i'^{(k)} / A_i \quad (32).$$

Ma poiché $b_i'^{(k)}$ è una stima via via più precisa della radiosità finale b_i , ora $\Delta b_i^{(k)}$ è una stima della quantità residua di radiosità che all'elemento i resta da irradiare dopo il passo k. A causa di questa radiosità residua, le immagini ottenute durante i primi passi di iterazione sono scure, ed a poco a poco diventano più chiare avvicinandosi all'immagine finale. Purtroppo, però, questa stima $b_i'^{(k)}$ tiene sì conto dell'energia residua $\Delta b_i^{(k)} = r_i^{(k)} / A_i$ immagazzinata nell'elemento i al passo k di iterazione, ma in modo impreciso: al termine del ciclo iterativo il residuo di energia immagazzinata non contribuisce solo alla radiosità dell'elemento i che la immagazzinava, ma anche a quella di tutti gli altri, in seguito alle ulteriori emissioni e ricezioni ed in proporzione alle loro riflettività ed ai fattori di forma. Occorre quindi una stima più accurata che tenga conto di questa distribuzione.

CORREZIONE AMBIENTALE DELLA LUMINOSITÀ NEL CORSO DELLE ITERAZIONI

Se l'algoritmo di radiosità è formulato in termini dell'emissione di energia (o radiosità) dagli elementi, invece che della loro ricezione di energia dagli altri elementi, e l'approssimazione iniziale della soluzione è il vettore nullo, allora la sola illuminazione dopo il primo ciclo di iterazione è quella proveniente direttamente dalle sorgenti di luce. Nelle iterazioni successive si considerano una riflessione, due riflessioni e così via. Chiaramente, nelle immagini ottenute dalle prime fasi di iterazione manca l'effetto di molteplici riflessioni successive, e quindi queste immagini sono scure.

Vogliamo usare la stima (32) per determinare un termine di correzione ambientale da aggiungere alla radiosità di tutte gli elementi per compensare la bassa luminosità delle immagini iniziali: grazie ad esso si possono ottenere approssimazioni iterative dell'immagine finale già buone dopo non molte iterazioni, e quindi si velocizza considerevolmente il metodo della radiosità. Come accennato prima, questo algoritmo è stato introdotto in [M.F.Cohen, S.E.Chen, J.R.Wallace e D.P.Greenberg, [A Progressive Refinement Approach to Fast Radiosity Image Generation](#), SIGGRAPH '89, in *Computer Graphics*, 23(3), luglio 1989].

Poiché questa correzione deve basarsi su una stima dell'effetto delle riflessioni successive, dobbiamo calcolarla stimando il coefficiente di riflettività media della scena, che si può definire come la media dei coefficienti di riflettività ρ_i degli elementi, pesata rispetto alla proporzione dell'area del singolo elemento rispetto all'area totale:

$$\rho_{media} = \sum_{i=1,...,n} \rho_i A_i / \sum_{i=1,...,N} A_i \quad (33)$$

Cosa si troverebbe se si cercasse una stima della luminosità ambientale residua già al passo iniziale di iterazione? A quel passo l'energia inviata all'ambiente per ogni elemento è data dal vettore e (la luce propria emessa dalle sorgenti). Allora, dopo la prima riflessione avremmo un contributo addizionale $\rho_{media}e$, dopo la seconda un ulteriore contributo $(\rho_{media})^2 e$, e così via: l'incremento di luminosità è quindi Re , dove il numero R è il fattore di interriflessione globale

$$R = 1 + \rho_{media} + (\rho_{media})^2 + (\rho_{media})^3 + \dots = 1/(1 - \rho_{media}) \quad (34)$$

Osserviamo che $0 < \rho_{media} < 1$, e quindi $R > 1$. L'incremento di luminosità dà quindi l'impressione falsa che l'energia non si conservi: ma in realtà qui stiamo misurando non l'energia ma la radiosità, ovvero la densità di potenza per unità di area: si parla quindi del flusso temporale di densità di energia. Questi effetti di maggiore luminosità sono dovuti a riflessioni successive della luce, che avvengono in istanti successivi del tempo: in ciascun istante l'energia totale rimane sempre quella iniziale, ma il fattore di correzione, che

amplifica indebitamente questa energia, avrebbe l'effetto di ridistribuirla creando una illuminazione di fondo più precisa, se applicassimo il ragionamento non alla prima iterazione ma via via a quelle successive. Vediamo come.

Una stima dell'energia totale H ancora non irradiata nell'ambiente dopo la k -sima iterazione (quindi dopo k interriflessioni) è la somma delle energie residue immagazzinate dagli elementi:

$$H = \sum_{i=1, \dots, n} r_i^{(k)} \quad (35)$$

Ma a noi serve la correzione ambientale delle radiosità, non delle energie. Perciò dobbiamo utilizzare il flusso per unità di area di questa energia totale non irradiata. L'area totale dell'ambiente è la somma A delle aree degli elementi:

$$A = \sum_{i=1, \dots, n} A_i .$$

Quindi il flusso per unità di area dell'energia totale non irradiata è

$$\Phi = RH/A \quad (36)$$

Questo flusso ambientale, effetto globale di tutte le interriflessioni successive, colpisce tutte gli elementi, e le illumina in proporzione al loro coefficiente di riflettività: pertanto al termine della k -sima iterazione la stima della radiosità con incremento ambientale è data da

$$b_i'^{(k)} = b_i^{(k)} + \rho_i \Phi \quad (37)$$

Al crescere di k , $r_i^{(k)}$ tende a 0 per ogni $i = 1, \dots, n$, e quindi H tende a 0. Pertanto Φ tende a 0, e l'effetto di correzione si attenua progressivamente col procedere delle iterazioni. Esso è intenso solo all'inizio, e permette di simulare con buona approssimazione quanto devono essere schiarite le immagini approssimate iniziali perchè la loro luminosità totale (ma ovviamente non la loro distribuzione di luce) sia accettabile.

ACCURATEZZA NEL CALCOLO DEI FATTORI DI FORMA, USO DEL RAY TRACING E MAPPA DI OCCLUSIONE

Nonostante l'utilizzo di uno [z-buffer](#) hardware renda il metodo di calcolo dei fattori di forma basato sulla proiezione sul semicubo un algoritmo estremamente efficiente per quanto riguarda il tempo di calcolo, la stessa cosa non si può dire per quanto riguarda l'accuratezza.

Ecco alcuni problemi:

1 – A causa del modo in cui funziona lo [z-buffer](#), ogni pixel del semicubo viene associato ad un solo elemento, anche nel caso in cui in quel pixel siano parzialmente visibili due o più elementi. Questo può portare a fenomeni di aliasing fra gli elementi che si sovrappongono quando proiettati su un lato del semicubo, ed alla scomparsa di parti di elementi o all'apparire di una alternanza a mosaico di due elementi lungo le linee sul buffer dove si intersecano due elementi diversi. Questi errori sono confinati ad aree grandi quanto un pixel, quindi piccole, ma queste aree sono piccole *solo per il calcolo dei fattori di forma*: la loro visibilità dal punto di vista dell'osservatore potrebbe essere elevata, ossia esse potrebbero coprire zone grandi nell'immagine resa, che sarebbero visti con effetti non corretti di illuminazione.

2 - Inoltre, si assume che il punto centrale di un elemento sia rappresentativo della ubicazione dell'intero elemento quando da esso si guardano gli altri (e viceversa). Nei casi in cui ciò non è vero l'elemento può essere suddiviso in [sottoelementi](#), ma questo si fa una volta per tutte, e quindi una volta sola: se il punto centrale non è adeguato quando osserviamo un altro elemento, allora procediamo a suddividere, ma da quel momento, se esso rimane inadeguato anche dopo tale suddivisione quando osserviamo un terzo elemento, non si può risuddividere ulteriormente perché ormai i sottoelementi sono stati già utilizzati nel calcolo, ed occorrerebbe ripetere il calcolo da capo con una sottostrutturazione ancora più fine.

3 – Infine osserviamo che, affinché l'intero procedimento basato sul semicubo dia un risultato corretto, gli elementi devono essere lontani fra di loro. La stima dei fattori di forma è approssimata per difetto per elementi vicini o adiacenti: sulle zone di contatto il calcolo dei fattori di forma precalcolati per i pixel dell'emicubo è sottostimato, perché è relativo al centro del pixel (dove i due elementi sono un po' più lontani), e non al punto di contatto che corrisponde invece di solito al bordo del pixel.

Il calcolo progressivo della radiosità è stato anche effettuato mediante l'impiego, nella valutazione dei fattori di forma, del Ray Tracing (non ricorsivo) al posto della proiezione sul semicubo [J.R.Wallace, K.A.Elmquist e E.A.Haines, [A Ray Tracing Algorithm for Progressive Radiosity](#), SIGGRAPH 89, in *Computer Graphics* 23(3), luglio 1989, pagg.

315-324]. Questo si fa nel modo seguente. Quando un elemento diffonde nell'ambiente la propria radiosità, per ogni vertice della scena vengono tracciati raggi verso l'elemento. Si suddivide l'elemento in un numero finito di sottoelementi, ognuno dei quali è il bersaglio di un raggio uscente dal vertice. Se il raggio non viene interrotto lungo il suo cammino, allora il sottoelemento dell'elemento è visibile e viene quindi calcolato, in modo analitico, il fattore di forma "infinitesimo" tra il vertice ed il sottoelemento, grazie alla solita formula

$$F_{di-j} = \int_{A_j} \frac{\cos \theta_i \cos \theta_j}{\pi r^2} H_{ij} dA_j, \text{ oppure lo stesso fattore di forma viene calcolato in modo}$$

numerico o statistico. A questo scopo si tracciano raggi a partire dal vertice, attraverso un viewport (o meglio un emisfero o un emicubo) suddiviso in pixel. I raggi sono centrati al centro dei vari pixel oppure distribuiti aleatoriamente (quale sia la migliore scelta di distribuzione di probabilità lo vedremo nel terzo volume di quest'opera, dove presenteremo metodi probabilistici di illuminazione globale). Essi vengono tracciati e si conta quanti di essi colpiscono il sottoelemento: la frazione di quelli che lo colpiscono misura la proporzione dell'angolo solido da esso sotteso rispetto a quel vertice, e quindi il fattore di forma infinitesimo del sottoelemento rispetto al vertice.

A questo punto il fattore di forma tra il vertice e l'intero elemento viene calcolato come somma dei fattori di forma infinitesimi dei sottoelementi visti dal vertice. In tal modo si ottengono direttamente i fattori di forma degli elementi visti dai vertici, il che è un vantaggio perché evita la fase di postprocessing nella quale le radiosità delle facce vengono trasformate in radiosità dei vertici: qui si ottengono direttamente le radiosità dei vertici (si veda l'articolo [M.F.Cohen e D.P.Greenberg, [The Hemi-Cube: A Radiosity Solution for Complex Environments](#), SIGGRAPH 85, in *Computer Graphics* 19(3), luglio 1985, pagg. 21-40] già citato [prima](#), alla fine della sezione "[L'equazione della radiosità](#)").

Si noti che il fattore di forma di un elemento visto da un punto è una misura dell'occlusione frapposta dall'elemento all'illuminazione del punto dovuta al resto dell'ambiente, ossia del suo livello di ombra. In questo senso, il calcolo del fattore di forma tramite Ray Tracing dei vari elementi rispetto ad un punto equivale alla *mappa di occlusione* che ora definiamo e spieghiamo, a costo di una breve digressione. Consideriamo un punto, ad esempio appartenente ad una faccia piana di un elemento, ed il semispazio ad esso frontale. La frazione di angolo solido coperta dagli altri elementi, misurata attraverso il numero di *hits* (intersezioni) di raggi proiettori uscenti dal punto, misura l'intensità di ombra a quel punto come conseguenza dell'occlusione della luce dovuta agli altri elementi. Essa si può reinterpretare nel modo seguente. Supponiamo dapprima che il punto stia su un elemento da illuminare con luce puramente ambientale, non riflettente né diffusiva secondo il modello di Lambert. Immaginiamo che il semispazio frontale sia racchiuso in un emisfero uniformemente illuminato e non riflettente, che funge da sfondo luminoso, ossia da sorgente di luce diffusa. Emaniamo, dal punto in oggetto, un gran numero di raggi proiettori, approssimativamente angolarmente equidistribuiti nell'emisfero, e tracciamoli nella scena finché colpiscono un altro elemento o l'emisfero. Ciascun raggio proiettore che non viene bloccato dagli altri elementi colpisce l'emisfero, e, nel consueto meccanismo del Ray Tracing, ogni *hit* fornisce un contributo di energia luminosa. La somma di questi contributi è l'illuminazione del punto causata dallo sfondo luminoso: se si preferisce, è la luce ambientale che illumina il punto, detratta però di quella bloccata dagli altri elementi, ossia attenuata a causa dell'ombra che essi proiettano. In tal modo la mappa di occlusione

definisce ombre, e quindi differisce dal modello banale della luce ambientale perché ha aspetti direzionali: tiene conto delle direzioni in cui il punto vede gli altri elementi.

Se invece si intende simulare l'illuminazione, e quindi il livello di ombra, di un punto che giace in un elemento [diffusivo secondo Lambert](#), è ancora possibile calcolare l'occlusione della scena mediante Ray Tracing. In questo caso dobbiamo immaginare che l'emisfero uniformemente luminoso si scomponga in tante piccole sorgenti di luce puntuali che emettono luce isotropicamente, tutte alla stessa intensità. In tal caso, ogni raggio proiettore, in base al modello di Lambert, colpisce l'emisfero in una delle piccole sorgenti di luce, e fornisce un contributo di illuminazione proporzionale al coseno dell'angolo di deviazione dal polo Nord: infatti questo è esattamente l'angolo fra la direzione della piccola sorgente di luce verso cui punta il raggio proiettore ed il versore normale al punto di emanazione, che è nient'altro che il versore verticale, ossia diretto verso il polo Nord. Pertanto basta, in questo caso, emanare raggi proiettori non più equidistribuiti, ma pesati secondo il coseno dell'angolo di deviazione dal polo Nord. Questo meccanismo pesa diversamente le sorgenti puntiformi di cui idealmente si compone l'emisfero. Vengono pesate di più le aree dell'emisfero frontali rispetto al punto di emanazione, cioè quelle in un intorno del polo Nord (dove la deviazione è zero ed il coseno vale 1), le quali inducono una illuminazione più intensa proprio perché frontali, in perfetto accordo con la legge empirica di Lambert. Queste aree vengono campionate maggiormente perché un maggior numero di raggi vengono emessi in tali direzioni, e quindi contribuiscono di più all'illuminazione del punto di emissione, naturalmente se esse non sono occluse.

Infine, se il punto giace in un elemento che oltre ad essere diffusivo è anche parzialmente speculare, e la scena è illuminata non isotropicamente, l'illuminazione e l'ombra si possono ancora calcolare come sopra. Supponiamo che la distribuzione di luci della scena sia tale che, ad esempio, ci sia una direzione dalla quale proviene principalmente la luce, mentre nelle altre direzioni l'emisfero è relativamente buio. Allora i raggi devono essere distribuiti in maniera da privilegiare la direzione di provenienza della luce: la distribuzione angolare dei raggi deve essere più fitta in tale direzione, e deve decrescere come il coseno dell'angolo di deviazione.

Appofittiamo di queste osservazioni per riformularle in un senso che diventerà cruciale quando, nel terzo volume di quest'opera, lo riconsidereremo nell'ambito dell'illuminazione globale. Abbiamo già osservato che il calcolo dei fattori di forma, ovvero della mappa di occlusione, tramite Ray Tracing richiede, come tipico del Ray Tracing, che vengano emessi molti raggi. C'è però un modo di evitare questo proliferare di raggi da emettere con una distribuzione che può essere anisotropa: invece di emettere molti raggi deterministicamente, ne emettiamo pochi ma stocasticamente, ovvero con una distribuzione di probabilità corrispondente alla distribuzione angolare voluta. Il senso di questa corrispondenza ed il modo in cui viene realizzata la distribuzione di probabilità vengono studiati nel terzo volume, nel capitolo sui metodi probabilistici per il calcolo di integrali e la loro applicazione all'illuminazione globale.

RIFLETTIVITA' BIDIREZIONALE

Il metodo di radiosità sviluppato fino a questo momento tratta solo la riflessione diffusa. In tal modo, la radiosità uscente da un elemento è influenzata dalla radiosità totale e non dalle direzioni attraverso cui viene ricevuta energia. L'articolo [*D.S.Immel, M.F.Cohen e D.P.Greenberg, [A Radiosity Method for Non-Diffuse Environments](#), SIGGRAPH 86, in *Computer Graphics* 20(4), agosto 1986, pagg. 133-142]* estende il meccanismo di radiosità per modellare anche la riflessione speculare. Invece di calcolare un singolo valore di radiosità per ogni elemento, viene calcolato un valore per ogni direzione. Più precisamente, si partiziona l'emisfero sopra l'elemento in un insieme finito di angoli solidi, ognuno dei quali individua una direzione per l'energia entrante o uscente (ovviamente, quindi, si limita l'attenzione ad un numero finito di direzioni: il problema viene discretizzato), e si calcola, in base ai modelli fisici o anche solo al modello euristico di Phong,, la percentuale della luce che entra dalla direzione entrante ed esce nella direzione di uscita (si vedano le sezioni sui [Modelli fisici di illuminazione](#) e sulla [Equazione integrale del rendering](#)). Pesando i contributi mediante questo *coefficiente di riflettività bidirezionale* ora possiamo calcolare la radiosità uscente in ogni direzione come somma dei contributi di radiosità entranti dalle varie direzioni (discretizzate, quindi in numero finito) ed emesse verso quella data direzione di uscita più la radiosità creata in quel dato elemento ed emessa verso quella direzione di uscita. L'immagine infine viene resa interpolando le intensità determinate per ogni vertice tramite media pesata delle radiosità direzionali uscenti dall'elemento a cui il vertice appartiene, interpolate a partire da quelle delle direzioni discretizzate più vicine alla direzione dal vertice all'osservatore. Questo tipo di approccio però porta ad una mole di calcoli estremamente onerosa.

METODI DI RENDERING A DUE PASSI CHE COMBINANO RADIOSITÀ E RAY TRACING RICORSIVO

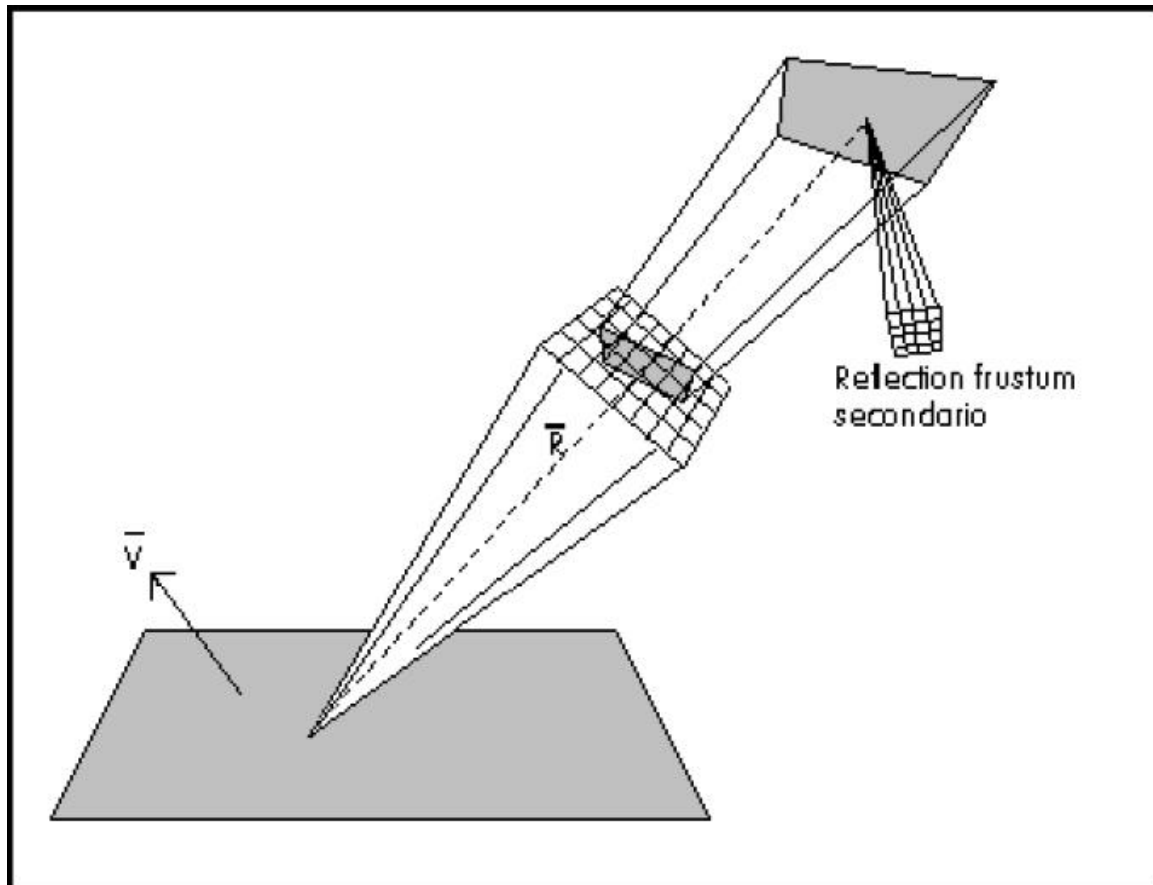
L'algoritmo di radiosità si adatta bene alla riflessione diffusa, ma non a quella speculare, poiché l'intensità della riflessione speculare dipende dall'angolo con il quale l'osservatore guarda la superficie, ma questo angolo non viene considerato nell'algoritmo di radiosità, che non è direzionale. D'altra parte il procedimento di [ray tracing](#) è ideale per il calcolo della riflessione speculare, ma non per la diffusione globale dell'ambiente. (Per maggiori dettagli, si veda la sezione sulla [Equazione integrale del rendering](#))

È quindi opportuno trovare il modo di combinare i due metodi in modo da sfruttare i vantaggi di entrambi. Purtroppo però non è sufficiente applicare entrambi i metodi e poi sommare i risultati, perché le componenti della luminosità dovute alla riflessione speculare contribuiscono anche alla illuminazione diffusa e viceversa.

Nell'articolo [[J.R.Wallace, M.F.Cohen e D.P.Greenberg, A Two-pass Solution to the Rendering Equation: a Synthesis of Ray Tracing and Radiosity Methods, SIGGRAPH '87, in Computer Graphics, 21\(4\), luglio 1987, pagg. 311-320](#)], viene sviluppato un approccio a due passi che combina il metodo di radiosità (indipendente dall'angolo di visuale) con quello di ray tracing, che invece ne dipende. Il primo passo, che non dipende dalla direzione di osservazione, consiste di un metodo ampliato di radiosità che tiene conto anche della riflessione speculare, perché amplia in maniera virtuale la scena, come se gli elementi riflettenti fossero finestre su un "mondo speculare". I fattori di forma da un elemento generico ad uno di questi elementi virtuali riflessi dall'altro lato di un elemento di "specchio" rappresentano il contributo alla diffusione globale da parte della riflessione speculare dello specchio. Purtroppo, se le superficie speculari sono ampie, questa fase aumenta quasi di un fattore due i fattori di forma da calcolare, e quindi rende quattro volte più lento il tempo di calcolo.

Nel secondo passo, dipendente dal punto di visuale, si considera ogni punto visibile di ogni elemento (compresi quelli virtuali). Del raggio proiettore proveniente a questo punto dall'osservatore si calcola la direzione di riflessione speculare, e si costruisce, anziché il raggio riflesso del ray tracing, una piramide di riflessione, cioè uscente in tale direzione e di apertura piccola, perché limitata all'angolo solido che include solo le deviazioni dalla direzione di riflessione speculare per le quali l'intensità riflessa è sufficientemente elevata (per determinare quanto elevata sia l'intensità in ciascuna direzione uscente si calcola il coefficiente di riflettività bidirezionale: ancora una volta, si vedano le sezioni sui [Modelli fisici di illuminazione](#) e sulla [Equazione integrale ricorsiva del rendering](#), nonché la sezione precedente). Sul piano perpendicolare all'asse della piramide (cioè alla direzione di riflessione speculare) si considera un piccolo rettangolo che viene usato per eseguire uno z-buffer di piccole dimensioni: così si determina quali altri elementi siano visibili attraverso ciascun pixel, e di ciascuno di essi si considera l'illuminazione diffusa calcolata nella prima fase tramite la radiosità estesa (la radiosità è stata riportata ai vertici della scena, ed ora la si riporta ai punti visti attraverso i centri dei pixel mediante interpolazione di Gouraud). Se c'è più di un elemento visibile attraverso la piramide, si calcola il contributo totale risultante tramite interpolazione (ossia combinazione pesata in termini alla proporzione delle aree visibili): si osservi che questa fase è analoga al calcolo dei fattori di forma differenziali. Inoltre si tiene conto se la piramide attraverso il pixel vede altri elementi speculari: per

ciascuno di essi ripetiamo il procedimento in maniera ricorsiva (per un numero prefissato di generazioni) emettendo un'altra piramide riflessa (esattamente come nel procedimento di generazione iterativa di raggi riflessi del ray tracing ricorsivo). Si totalizzano i contributi di illuminazione ottenuti, come spiegato prima, da ognuna delle piramidi così generate, ed in tal modo si ottiene il contributo della riflessione speculare al vertice della piramide iniziale. Anche il calcolo di questa fase, ovviamente, è piuttosto oneroso.



Nella figura qui sopra è mostrata una piramide di riflessione. Riassumiamo il procedimento. Quale sia per ogni pixel l'elemento speculare visibile si calcola col metodo di z-buffer; si utilizza l'interpolazione di Gouraud per calcolare per quel pixel le intensità di riflessione diffusa del primo passo. Se l'elemento visibile ad un dato pixel di questo z-buffer è un elemento che dà luogo a riflessione speculare, allora si traccia il raggio dal vertice della piramide attraverso il centro del pixel e si determina l'angolo di riflessione speculare al punto in cui colpisce l'elemento: continuando iterativamente così si disegnano nuove piramidi in corrispondenza di ogni intersezione. Attraverso una media dei valori calcolati per ogni pixel dello z-buffer di ciascuna piramide si trova l'intensità della luce riflessa sull'elemento di partenza.

Questo metodo riesce a combinare la radiosità col ray tracing ricorsivo, ma porta ad una proliferazione dei fattori di forma, il cui calcolo è lento. Abbiamo già osservato che, se tutti gli elementi riflettono luce anche specularmente, contando anche gli elementi virtuali il loro numero si raddoppia ed il numero dei fattori di forma si quadruplica. Una variante più efficiente, introdotta in [F.Sillion, C.Puech, [A General Two-Pass Method Integrating](#)

[Specular and Diffuse Reflection](#), SIGGRAPH '89, in *Computer Graphics*, 23(3), luglio 1989] calcola fattori di forma estesi, i quali sono in grado di tener conto di un numero arbitrario di riflessioni speculari o rifrazioni. Questi fattori di forma estesi si calcolano come sopra ma mediante Ray Tracing ricorsivo fin dal primo passo, invece che mediante emicubi e z-buffer come nel meccanismo usuale della radiosità, oppure mediante Ray Tracing non ricorsivo come nel caso della [mappa di occlusione](#) vista prima.

ESEMPIO 1: DISTRIBUZIONE DELLA LUCE ENTRO UNA TENDA PIRAMIDALE

La scena consiste dell'interno di una tenda, cioè una stanza a forma di piramide equilatera, quindi con quattro facce identiche a forma di triangolo equilatero. La faccia di base emette luce con intensità 1 (possiamo pensare che ci sia un fuoco acceso al centro della base: in tal caso la luce viene emessa al centro, ma nella presente approssimazione, peraltro alquanto rudimentale, l'intero triangolo di base consiste di un unico elemento, che indichiamo col numero 1). Le altre facce non emettono luce propria, solo riflessa. La base, cioè il pavimento, è di terra scura, e l'approssimiamo come una superficie non riflettente, cioè $\rho_1 = 0$. Le pareti sono grigio chiaro, e quindi poniamo $\rho_i = \frac{1}{2}$ per $i=2, 3, 4$. Per comodità scriviamo ρ invece di $\frac{1}{2}$. Quindi il vettore dell'energia propria emessa è $\mathbf{e} = (1, 0, 0, 0)$, mentre il vettore della riflettività è $\mathbf{p} = (0, 1/2, 1/2, 1/2)$.

Però in questo caso gli elementi sono piane, e quindi i fattori di forma diagonali, F_{i-i} , sono tutti nulli. D'altra parte, la tenda è completamente simmetrica: ogni faccia copre, vista da ogni altra faccia, lo stesso angolo solido. Perciò, per i diverso da j , i fattori di forma F_{i-j} sono tutti uguali, cioè c'è un unico fattore di forma non nullo che per semplicità chiamiamo F . (Si osservi che, poiché le aree di tutte gli elementies sono uguali, questo risultato conferma anche l'equazione di reciprocità data dalla reversibilità dei percorsi ottici, $A_i F_{ij} = A_j F_{ji}$: questa e le prossime equazioni sono state introdotte nella sezione [L'equazione della radiosità](#). Ora, dall'equazione della conservazione dell'energia (o, se si preferisce, di conservazione dell'angolo solido totale) $\sum_j F_{ij} = 1$, si ottiene $F=1/3$. Pertanto il sistema lineare della radiosità diventa:

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ -\rho F & 1 & -\rho F & -\rho F \\ -\rho F & -\rho F & 1 & -\rho F \\ -\rho F & -\rho F & -\rho F & 1 \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

Procediamo a risolvere il sistema con metodi di rilassamento. Per poter applicare il metodo iterativo formulato in (6') della sezione [Analisi del metodo di Gauss-Seidel](#) dobbiamo spezzare M come $M = N - P$, con N matrice triangolare inferiore (o meglio non superiore: la diagonale di N coincide con quella di M) e P triangolare (strettamente) superiore. Quindi:

$$N = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -\rho F & 1 & 0 & 0 \\ -\rho F & -\rho F & 1 & 0 \\ -\rho F & -\rho F & -\rho F & 1 \end{pmatrix}$$

$$P = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & \rho F & \rho F \\ 0 & 0 & 0 & \rho F \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

1. Metodo di Jacobi

Per primo, consideriamo il metodo di Jacobi, introdotto nella sezione [Soluzione con metodi iterativi: rilassamento di Jacobi](#). In questo metodo, il ciclo di iterazione (equazione (3) della succitata sezione) viene eseguito su tutte le componenti dei vettori, quindi su tutte le n equazioni del sistema. Il vettore \mathbf{e} vale $(1,0,0,0)$. Per semplicità di calcolo conviene partire con l'approssimante iniziale $\mathbf{b}^{(0)} = (0,0,0,0)$. Si ottiene, da (6'):

$$\mathbf{b}^{(k+1)} = N^{-1}P\mathbf{b}^{(k)} + N^{-1}\mathbf{e} \quad (38)$$

Una matrice triangolare inferiore a dimensione 4, con 1 sulla diagonale, cioè del tipo

$$J = \begin{pmatrix} 1 & 0 & 0 & 0 \\ x & 1 & 0 & 0 \\ y & z & 1 & 0 \\ u & v & w & 1 \end{pmatrix}$$

ha per inversa (esercizio!):

$$J^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -x & 1 & 0 & 0 \\ xz - y & -z & 1 & 0 \\ (y - xz)w + xv - u & zw - v & -w & 1 \end{pmatrix}$$

Perciò

$$N^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ \rho F & 1 & 0 & 0 \\ \rho F + (\rho F)^2 & \rho F & 1 & 0 \\ \rho F + 2(\rho F)^2 + (\rho F)^3 & \rho F + (\rho F)^2 & \rho F & 1 \end{pmatrix}$$

Ora possiamo trovare gli approssimanti dati dalle prime iterazioni del metodo di Jacobi, cioè del rilassamento eseguito ricalcolando ad ogni ciclo iterativo tutte le n equazioni del sistema, ciascuna applicata ai dati provenienti dal ciclo precedente. Si trova, da (38):

$$\mathbf{b}^{(1)} = N^{-1}\mathbf{e}$$

$$\mathbf{b}^{(2)} = N^{-1}P\mathbf{b}^{(1)} + N^{-1}\mathbf{e} = N^{-1}P\mathbf{b}^{(1)} + \mathbf{b}^{(1)} = (N^{-1}P + I)\mathbf{b}^{(1)}$$

$$\mathbf{b}^{(3)} = N^{-1}P\mathbf{b}^{(2)} + N^{-1}\mathbf{e} = N^{-1}P\mathbf{b}^{(2)} + \mathbf{b}^{(1)}$$

Poiché il vettore e è il primo vettore della base canonica, il risultato dell'applicare ad e una matrice è il vettore dato dalla prima colonna della matrice. Pertanto:

$$b^{(1)} = \begin{pmatrix} 1 \\ \rho F \\ \rho F + (\rho F)^2 \\ \rho F + 2(\rho F)^2 + (\rho F)^3 \end{pmatrix}$$

$$b^{(2)} = (N^{-1}P + I) \begin{pmatrix} 1 \\ \rho F \\ \rho F + (\rho F)^2 \\ \rho F + 2(\rho F)^2 + (\rho F)^3 \end{pmatrix} = \begin{pmatrix} 1 \\ \rho F + 2(\rho F)^2 + 3(\rho F)^3 + (\rho F)^4 \\ \rho F + 2(\rho F)^2 + 4(\rho F)^3 + 4(\rho F)^4 + (\rho F)^5 \\ \rho F + 2(\rho F)^2 + 4(\rho F)^3 + 7(\rho F)^4 + 5(\rho F)^5 + (\rho F)^6 \end{pmatrix}$$

Nel nostro caso, $\rho=1/2$ e $F=1/3$, quindi $\rho F=1/6$. Se, per semplicità di calcolo, limitiamo la precisione alla prima cifra decimale, possiamo trascurare, in prima approssimazione, le potenze $(\rho F)^n$ con n maggiore di 2, perché $(\rho F)^2 = 0.02778$ e $(\rho F)^3 = 0.00077$ (si potrebbe trascurare già $(\rho F)^2$ se non fosse per il fatto che questo fattore compare moltiplicato per 2 nella terza componente e nella quarta componente del vettore $b^{(2)}$). Quindi, l'approssimazione con la prima cifra decimale esatta di $b^{(1)}$ è $b^{(1)} = (1, 0.17, 0.19, 0.23)$, mentre $b^{(2)} = (1, 0.24, 0.24, 0.25)$. Per $b^{(3)}$ si ottiene, ponendo per semplicità $t = \rho F$:

$$b^{(3)} = N^{-1}P \begin{pmatrix} 1 \\ t + 2t^2 + 3t^3 + t^4 \\ t + 2t^2 + 4t^3 + 4t^4 + t^5 \\ t + 2t^2 + 4t^3 + 7t^4 + 5t^5 + t^6 \end{pmatrix} + \begin{pmatrix} 1 \\ t \\ t + t^2 \\ t + 2t^2 + t^3 \end{pmatrix}$$

$$= \begin{pmatrix} 1 \\ t + 2t^2 + 4t^3 + 8t^4 + 11t^5 + 6t^6 + t^7 \\ t + 2t^2 + 4t^3 + 8t^4 + 15t^5 + 16t^6 + 7t^7 + t^8 \\ t + 2t^2 + 4t^3 + 8t^4 + 16t^5 + 26t^6 + 22t^7 + 8t^8 + t^9 \end{pmatrix} \approx \begin{pmatrix} 1 \\ 0.25 \\ 0.25 \\ 0.25 \end{pmatrix}$$

Si osservi che le componenti dei vettori $b^{(2)}$ e $b^{(3)}$ sono termini polinomiali in t nei quali i monomi di grado più basso si stabilizzano, cioè mantengono gli stessi coefficienti (lasciamo al lettore, come esercizio, dimostrare che il coefficiente del monomio t^n in tutte le componenti delle approssimazioni successive si stabilizza sul valore 2^{n-1}). Poiché alla precisione richiesta solo questi termini contano, per n maggiore di 2 le approssimazioni iterative $b^{(n)}$ sono costanti: abbiamo raggiunto l'equilibrio, cioè la soluzione $b = (1, 0.25, 0.25, 0.25)$. Si osservi che la prima componente, che corrisponde al pavimento della tenda piramidale, vale sempre 1 perché, pur non riflettendo luce, emette luce propria con potenza 1 per via della fiamma al suo centro. Osserviamo anche che all'equilibrio le tre pareti della tenda ricevono la stessa quantità di illuminazione: anche questo era previsto, per via della simmetria della scena fra le tre pareti.

E' evidente da questi calcoli che le iterazioni del metodo di Jacobi, in ciascuna delle quali si debbono ricalcolare tutte le equazioni del sistema, sono laboriose.

Presentiamo qui un notebook di [Mathematica](#) (software di manipolazione simbolica e numerica) che esegue i calcoli precedenti in maniera automatica:

```
P = {{0, 0, 0, 0}, {0, 0, t, t}, {0, 0, 0, t}, {0, 0, 0, 0}}
NN = {{1, 0, 0, 0}, {-t, 1, 0, 0}, {-t, -t, 1, 0}, {-t, -t, -t, 1}}
e = {1, 0, 0, 0}
Ninv = Inverse[NN]; TraditionalForm[Ninv]
b = {0, 0, 0, 0}
For[i=1, i<5, i++, Expand[b = Ninv . P . b + Ninv . e]; N[b] /. t -> 1/6]
```

Il lettore che non disponga di Mathematica può facilmente convertire al linguaggio C il precedente notebook e quelli che presenteremo in seguito; però poi, per gli integrali numerici necessari per i fattori di forma, deve usare le sofisticate librerie matematiche disponibili in C.

2. Metodo di Gauss-Seidel

Trattiamo ora lo stesso esempio con il metodo di Gauss-Seidel, introdotto nella sezione [Rilassamento di Gauss-Seidel applicato al sistema lineare della radiosità](#). Il ciclo di iterazione (equazione (3) della sezione [Soluzione con metodi iterativi: rilassamento di Jacobi](#)) viene eseguito su tutte le componenti dei vettori, quindi su tutte le n equazioni del sistema, però ora in ciascuna equazione, diciamo la j -sima ($j=1,\dots,n$), vengono utilizzate le componenti già aggiornate del vettore b : in altre parole, nell'iterazione $k+1$ -sima al secondo membro dell'equazione j -sima si utilizzano le componenti $b_i^{(k+1)}$ invece di $b_i^{(k)}$, per $i=1,\dots,j-1$. Partiamo di nuovo con l'approssimante iniziale $b^{(0)} = (0,0,0,0)$. Usiamo ancora (38) dell'[Esempio 1](#), ma con le sostituzioni appena indicate. Ecco i risultati:

$$\mathbf{b}^{(1)} = (1, 0.167, 0.194, 0.233)$$

$$\mathbf{b}^{(2)} = (1, 0.238, 0.245, 0.249)$$

$$\mathbf{b}^{(3)} = (1, 0.249, 0.25, 0.25)$$

$$\mathbf{b}^{(4)} = (1, 0.25, 0.25, 0.25)$$

Il numero di iterazioni necessarie per arrivare al punto di equilibrio è lo stesso del metodo di Jacobi illustrato nell'[Esempio 1](#) della tenda piramidale, ma ora ciascun approssimante è più vicino al risultato finale di quanto non accadesse prima.

Ecco il corrispondente notebook in Mathematica:

```
P = {{0, 0, 0, 0}, {0, 0, t, t}, {0, 0, 0, t}, {0, 0, 0, 0}}
NN = {{1, 0, 0, 0}, {-t, 1, 0, 0}, {-t, -t, 1, 0}, {-t, -t, -t, 1}}
e = {1, 0, 0, 0}
TraditionalForm[Ninv = Inverse[NN]]
b = {0, 0, 0, 0}
For[k = 1, k < 5, k++,
For[i = 1, i < 5, i++,
Expand[b[[i]] = (Ninv.P)[[i]]. b + Ninv[[i]]. e]; N[b] /. t -> 1/6 ]
```

3. Metodo di Southwell

Trattiamo infine lo stesso esempio con il metodo di Southwell, introdotto nella sezione [Rilassamento di Southwell applicato al sistema lineare della radiosità](#). Il ciclo di iterazione dovrebbe ora riguardare la matrice dell'energia K invece di quella della radiosità M, ma poiché le aree degli elementi sono tutte uguali le due matrici coincidono. Quindi il ciclo iterativo rimane lo stesso (equazione (3) nella sezione [Soluzione con metodi iterativi: rilassamento di Jacobi](#)), ma viene eseguito ora su una sola equazione del sistema, quella con il resto maggiore. Il vettore dei resti alla k-sima iterazione è proporzionale a $r^{(k)} = e - M b^{(k)}$, grazie a (13') della sezione [Rilassamento di Southwell applicato al sistema lineare della radiosità](#). Quando il vettore dei resti diventa nullo abbiamo raggiunto l'equilibrio. Per ciascuna iterazione calcoliamo il vettore dei resti, determiniamo l'indice (diciamo j) della sua componente più grande (in valore assoluto) e calcoliamo l'iterazione successiva svolgendo solo la j-sima equazione del sistema lineare della radiosità. Con i metodi precedentemente illustrati ci volevano da tre a quattro iterazioni, risolvendo 4 equazioni per ogni iterazione. Quindi ora potremmo aspettarci di dover svolgere da 12 a 16 iterazioni. Ma poiché ogni volta trattiamo proprio la componente con l'errore maggiore, il metodo è più efficiente, e bastano solo sette iterazioni. Forniamo i risultati con un numero maggiore di cifre decimali rispetto a prima per evidenziare la progressione dell'iterazione. Si noti, per verifica di quanto appena detto, che in ogni iterazione cambia una sola componente (mai la prima, perché quella ha errore zero, in quanto la corrispondente elemento ha coefficiente di riflettività zero).

$$b^{(1)} = (1, 0, 0, 0.226852)$$

$$b^{(2)} = (1, 0, 0.238555, 0.226852)$$

$$b^{(3)} = (1, 0.244234, 0.238555, 0.226852)$$

$$b^{(4)} = (1, 0.244234, 0.238555, 0.248236)$$

$$b^{(5)} = (1, 0.249596, 0.238555, 0.248236)$$

$$b^{(6)} = (1, 0.249596, 0.238555, 0.249872)$$

$$b^{(7)} = (1, 0.249596, 0.249957, 0.249872)$$

cioè, a due cifre decimali esatte (in realtà tre!), $b^{(7)} = (1, 0.25, 0.25, 0.25)$. Ecco il corrispondente notebook di [Mathematica](#). Si noti che tutte gli elementi hanno la stessa area A, quindi, a meno della moltiplicazione per A, il vettore resto alla k-sima iterazione coincide con $e - Mb^{(k)}$.

```
P = {{0, 0, 0, 0}, {0, 0, t, t}, {0, 0, 0, t}, {0, 0, 0, 0}}
NN = {{1, 0, 0, 0}, {-t, 1, 0, 0}, {-t, -t, 1, 0}, {-t, -t, -t, 1}}
e = {1, 0, 0, 0}
MatrixForm[M = NN - P]
TraditionalForm[Ninv = Inverse[NN]]
b = {0, 0, 0, 0}
```

```

maxr = 1.; While[ (maxr > 0.001),
r = N[e - M.b] /. t -> 1/6;
maxr = Max[r[[1]], r[[2]], r[[3]], r[[4]]];
For [i = 1, i < 5, i++, If [(r[[i]] == maxr), i0 = i, Continue]
];
b[[i0]] = N[(Ninv .P)[[i0]]. b + Ninv[[i0]] .e] /. t -> 1/6;
Print[TraditionalForm[N[b]]]
]

```

ESEMPIO 2: DISTRIBUZIONE DELLA LUCE ENTRO UNA STANZA CUBICA

La scena ora consiste dell'interno di una stanza a forma di cubo, quindi con sei facce identiche quadrate. La faccia alta ha una lampada che emette luce con intensità 1. Le altre facce non emettono luce propria, solo riflessa. Non ci sono finestre. La base, cioè il pavimento, è più scura degli altri muri: fissiamo $\rho_1 = 1/4$. Le pareti ed il soffitto sono grigio chiaro, e quindi poniamo $\rho_i = 1/2$ per $i=2, 3, 4, 5, 6$. Anche in questo caso ci limitiamo all'approssimazione rudimentale nella quale ciascuna parete forma un unico elemento (quindi non noteremo gradazioni di luce sulle pareti). Per comodità scriviamo ρ_+ invece di $1/2$ e ρ_- invece di $1/4$. Quindi il vettore dell'energia propria emessa è $\mathbf{e} = (0,0,0,0,0,1)$, mentre il vettore della riflettività è $\mathbf{\rho} = (1/4, 1/2, 1/2, 1/2, 1/2, 1/2)$. Nel seguito scriveremo \mathbf{e} invece di \mathbf{e} .

Anche in questo caso gli elementi sono piani, e quindi i fattori di forma diagonali, F_{i-i} , sono tutti nulli. D'altra parte, la stanza è completamente simmetrica. Quindi ogni faccia copre, vista da ogni altra faccia ad essa laterale, lo stesso angolo solido. Inoltre ogni faccia copre lo stesso angolo solido (ma diverso da quello di prima) quando è vista dalla faccia diametralmente opposta. Perciò ci sono solo due fattori di forma non nulli: quello laterale, che per semplicità chiamiamo F_l , e quello frontale, F_f . In particolare, $F_{ij} = F_f$ per $(i,j)=(1,6), (2,4), (3,5)$ ed i loro trasposti, $F_{ij} = F_l$ per $(i,j)=(1,2), (1,3), (1,4), (1,5), (6,2), (6,3), (6,4), (6,5), (2,3), (2,5), (3,4), (4,5)$ e tutti i loro trasposti. Si noti che ciascuna faccia ha quattro facce laterali ed una frontale. Quindi, dall'equazione della conservazione dell'energia (o dell'angolo solido totale) $\sum_j F_{ij} = 1$, che abbiamo introdotto nella sezione [L'equazione della radiosità](#), si ottiene $F_f + 4 F_l = 1$. Il valore di F_f si può calcolare grazie alla formula dimostrata nella sezione [Calcolo dei fattori di forma](#):

$$F_{i-j} = \frac{1}{A_i} \int_{A_i} \int_{A_j} \frac{\cos \theta_i \cos \theta_j}{\pi r^2} H_{ij} dA_j dA_i$$

Qui non ci sono ostruzioni dentro la stanza, quindi H vale costantemente 1, e le facce antistanti sono parallele, quindi i due angoli θ_i e θ_j sono uguali. Chiameremo quest'angolo θ . L'integrale si può approssimare numericamente con il metodo illustrato nella sezione [Calcolo dei fattori di forma](#), legato al metodo di [z-buffer](#). Ma in questo caso l'integrale si può anche calcolare analiticamente come segue. Per simmetria, possiamo supporre che $i=1$ e $j=6$, cioè che la faccia i sia quella inferiore e la faccia j quella superiore. Per l'invarianza sotto cambiamento di scala degli angoli solidi sottesi da facce antistanti, possiamo supporre che il cubo abbia lato 1, quindi tutte le facce abbiano area 1. Possiamo anche supporre che il cubo abbia le facce parallele ai piani coordinati e giaccia nell'ottante positivo, con un vertice nell'origine. Allora la sua faccia inferiore ha coordinate $(x,y,0)$ per x e y che variano fra 0 e 1. Analogamente, la faccia superiore ha coordinate $(x',y',1)$ per x' e y' che variano fra 0 e 1.

Perciò, a meno del fattore π al denominatore, l'integrando ha la forma seguente. Chiamiamo \mathbf{p} il vettore che congiunge un punto generico della faccia inferiore ad uno della

faccia superiore: $p = (x', y', 1) - (x, y, 0) = (x' - x, y' - y, 1)$. Osserviamo che la terza coordinata di p vale costantemente 1. Ora normalizziamo p , ottenendo $q = p/||p||$. Il valore di $\cos \theta$ è esattamente il prodotto scalare fra il versore q ed il versore perpendicolare alla due facce, cioè il versore dell'asse z , $(0, 0, 1)$. Pertanto $\cos \theta$ è nient'altro che la terza componente del versore q : cioè $\cos \theta = 1/||p||$. D'altra parte, la distanza r fra $(x, y, 0)$ e $(x', y', 1)$ è proprio $||p||$. Quindi l'integrando è $1/(\pi ||p||^4)$. (Questi risultati sono stati già osservati nella sezione [Calcolo dei fattori di forma](#)).

Questo integrando è il reciproco di un polinomio di secondo grado nelle quattro variabili x, y, x', y' , e quindi si può calcolare analiticamente in termini di funzioni arcoseno e logaritmo. Il calcolo è laborioso, ed è stato svolto in un Esempio precedente ([Calcolo del fattore di forma fra due pareti adiacenti di una stanza cubica](#)). Il risultato è:

$$F_f = 0.2$$

Chi vuole può riverificare il risultato analitico tramite approssimazione numerica. Ecco il listato di un notebook di Mathematica che svolge il calcolo:

```
pointdown[x, y] = {x, y, 0};
pointup[x', y'] = {x', y', 1};
displacement[x, y, x', y'] = pointup[x', y'] - pointdown[x, y];
NormSquare[x, y, x', y'] = 0;
For[i = 1, i < 4, i++,
  NormSquare[x, y, x', y'] = NormSquare[x, y, x', y'] +=
  displacement[x, y, x', y'][[i]]^2];
Norm[x, y, x', y'] = NormSquare[x, y, x', y']^(1/2);
Integrand[x, y, x', y'] = 1/ NormSquare[x, y, x', y']^2;
N[F = (1/Pi) Integrate[Integrand[x, y, x', y'], {x, 0, 1}, {y, 0, 1}, {x', 0, 1}, {y', 0, 1} ]
```

Pertanto, dall'equazione della conservazione dell'energia $F_f + 4 F_l = 1$ (introdotta nella sezione [L'equazione della radiosità](#)), segue $F_l = F_f = 0.2$. Il fatto che $F_l = F_f$ è insolito, ed è vero solo per le prime due cifre decimali e solo perché la stanza ha forma cubica. In seguito ([Esempio 3](#)) vedremo esempi di stanze di forma rettangolare dove questo non accade. Nel seguito indicheremo con F il valore comune $F_l = F_f$.

Quindi la matrice M del sistema lineare della radiosità è la seguente:

$$\begin{pmatrix} 1 & -\rho_- F & -\rho_- F & -\rho_- F & -\rho_- F & -\rho_- F \\ -\rho_+ F & 1 & -\rho_+ F & -\rho_+ F & -\rho_+ F & -\rho_+ F \\ -\rho_+ F & -\rho_+ F & 1 & -\rho_+ F & -\rho_+ F & -\rho_+ F \\ -\rho_+ F & -\rho_+ F & -\rho_+ F & 1 & -\rho_+ F & -\rho_+ F \\ -\rho_+ F & -\rho_+ F & -\rho_+ F & -\rho_+ F & 1 & -\rho_+ F \\ -\rho_+ F & -\rho_+ F & -\rho_+ F & -\rho_+ F & -\rho_+ F & 1 \end{pmatrix} \begin{pmatrix} B_1 \\ B_2 \\ B_3 \\ B_4 \\ B_5 \\ B_6 \end{pmatrix} = \begin{pmatrix} E_1 \\ E_2 \\ E_3 \\ E_4 \\ E_5 \\ E_6 \end{pmatrix}$$

cioè

$$\begin{pmatrix} 1 & -0.05 & -0.05 & -0.05 & -0.05 & -0.05 \\ -0.1 & 1 & -0.1 & -0.1 & -0.1 & -0.1 \\ -0.1 & -0.1 & 1 & -0.1 & -0.1 & -0.1 \\ -0.1 & -0.1 & -0.1 & 1 & -0.1 & -0.1 \\ -0.1 & -0.1 & -0.1 & -0.1 & 1 & -0.1 \\ -0.1 & -0.1 & -0.1 & -0.1 & -0.1 & 1 \end{pmatrix} \begin{pmatrix} B_1 \\ B_2 \\ B_3 \\ B_4 \\ B_5 \\ B_6 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

Applichiamo anche questa volta il metodo iterativo formulato in (6') della sezione [Analisi del metodo di Gauss-Seidel](#), spezzando M come $M = N - P$, con N matrice triangolare inferiore (o meglio non superiore) e P triangolare (strettamente) superiore. Abbiamo:

$$N = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ -0.1 & 1 & 0 & 0 & 0 & 0 \\ -0.1 & -0.1 & 1 & 0 & 0 & 0 \\ -0.1 & -0.1 & -0.1 & 1 & 0 & 0 \\ -0.1 & -0.1 & -0.1 & -0.1 & 1 & 0 \\ -0.1 & -0.1 & -0.1 & -0.1 & -0.1 & 1 \end{pmatrix}$$

$$P = \begin{pmatrix} 0 & 0.05 & 0.05 & 0.05 & 0.05 & 0.05 \\ 0 & 0 & 0.1 & 0.1 & 0.1 & 0.1 \\ 0 & 0 & 0 & 0.1 & 0.1 & 0.1 \\ 0 & 0 & 0 & 0 & 0.1 & 0.1 \\ 0 & 0 & 0 & 0 & 0 & 0.1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Calcoliamo l'inversa di N (possiamo usare il metodo di Gauss, o anche il metodo diretto usato prima per la tenda piramidale ([Esempio 1](#))). Per semplicità ricaviamo il risultato dalla stesso notebook di Mathematica usato in quella circostanza per verifica. Si ottiene:

$$N^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0.1 & 1 & 0 & 0 & 0 & 0 \\ 0.11 & 0.1 & 1 & 0 & 0 & 0 \\ 0.121 & 0.11 & 0.1 & 1 & 0 & 0 \\ 0.1331 & 0.121 & 0.11 & 0.1 & 1 & 0 \\ 0.14641 & 0.1331 & 0.121 & 0.11 & 0.1 & 1 \end{pmatrix}$$

1. Metodo di Jacobi

Ora gli approssimanti alle varie iterazioni successive si trovano come nel caso della tenda

piramidale dell'**Esempio 1**. Ecco i risultati col metodo di Jacobi:

$$b^{(1)} = (0.05, 0.105, 0.116, 0.127, 0.140, 1.053)$$

$$b^{(2)} = (0.077, 0.151, 0.155, 0.158, 0.159, 1.070)$$

$$b^{(3)} = (0.085, 0.163, 0.164, 0.164, 0.164, 1.074)$$

$$b^{(4)} = (0.086, 0.165, 0.165, 0.166, 0.166, 1.075)$$

$$b^{(5)} = (0.087, 0.166, 0.166, 0.166, 0.166, 1.075)$$

$$b^{(6)} = (0.087, 0.166, 0.166, 0.166, 0.166, 1.075)$$

Alla quinta iterazione l'approssimante ha raggiunto il punto di equilibrio (limitatamente alle prime tre cifre decimali: se si richiede una precisione maggiore allora sono necessarie più iterazioni).

Per completezza, riportiamo qui il notebook di Mathematica che abbiamo utilizzato:

```
t=0.1
P = {{0, t/2, t/2, t/2, t/2, t/2}, {0, 0, t, t, t, t}, {0, 0, 0,
t, t, t}, {0, 0, 0, 0, t, t}, {0, 0, 0, 0, 0, t}, {0, 0, 0, 0, 0,
0}}
NN = {{1, 0, 0, 0, 0, 0}, {-t, 1, 0, 0, 0, 0}, {-t, -t, 1, 0, 0,
0}, {-t, -t, -t, 1, 0, 0}, {-t, -t, -t, -t, 1, 0}, {-t, -t, -t, -t,
-t, 1}}
e = {0, 0, 0, 0, 0, 1}
Ninv = Inverse[NN]
TraditionalForm[Ninv]
b = {0, 0, 0, 0, 0, 0}
For[i = 0, i < 7, i++, N[Expand[b = Ninv . P . b + Ninv . e]];
Print[b] ]
```

2. Metodo di Gauss-Seidel

Ed ora i risultati col metodo di Gauss-Seidel (scriviamo in rosso le cifre che sono cambiate rispetto al metodo precedente):

$$b^{(1)} = (0.05, 0.105, 0.116, 0.130, 0.146, 1.068)$$

$$b^{(2)} = (0.078, 0.154, 0.158, 0.161, 0.164, 1.074)$$

$$b^{(3)} = (0.086, 0.164, 0.165, 0.165, 0.166, 1.075)$$

$$b^{(4)} = (0.087, 0.166, 0.166, 0.166, 0.166, 1.075)$$

$$b^{(5)} = (0.087, 0.166, 0.166, 0.166, 0.166, 1.075)$$

In questo esempio, il metodo di Gauss-Seidel impiega una iterazione in meno rispetto al metodo di Jacobi per raggiungere l'equilibrio, ed inoltre ad ogni iterazione i suoi approssimanti sono più vicini al risultato finale: se si richiedono tre sole cifre decimali, questa differenza è piuttosto evidente quando si è ancora lontani dall'equilibrio, cioè nelle prime tre iterazioni, soprattutto la seconda. Soprattutto in questo senso il metodo di Gauss-Seidel è più veloce.

Di nuovo, riportiamo il notebook di Mathematica che abbiamo utilizzato:

```
t=0.1
P={ {0,t/2,t/2,t/2,t/2,t/2}, {0,0,t,t,t,t}, {0,0,0,t,t,t}, {0,0,0,0,t,t}, {0,0,0,0,0,t}, {0,0,0,0,0,0} }
NN={ {1,0,0,0,0,0}, {-t,1,0,0,0,0}, {-t,-t,1,0,0,0}, {-t,-t,-t,1,0,0}, {-t,-t,-t,-t,1,0}, {-t,-t,-t,-t,-t,1} }
e={0,0,0,0,0,1}
Ninv=Inverse[NN]
b={0,0,0,0,0,0}
For[k=1,k<10,k++,
  For[i=1,i<7,i++,
    Expand[b[[i]]=(Ninv.P)[[i]].b+Ninv[[i]].e];
  Print[N[b]] ]
```

3. Metodo di Southwell

Infine, presentiamo i risultati ottenuti con il metodo di Southwell, secondo le linee esposte nella sezione [Metodo del rilassamento di Southwell applicato al sistema lineare della radiosità](#). La matrice dell'energia K, introdotta nella sottosezione [Rilassamento di Southwell](#), coincide con M perché le aree di tutti gli elementi sono uguali; inoltre valgono 1, quindi i vettori β della energia e b della radiosità coincidono. I risultati si ottengono rilassando una sol'elemento (invece che tutte e sei) per ogni ciclo di iterazione. (Si tratta dell'elemento col valore massimo del vettore resto $r^{(k)}$, il quale, grazie a (13'), vale $r^{(k)} = e - M b^{(k)}$).

Quindi mediamente potremmo aspettarci sei volte più iterazioni, ma, poiché ogni volta rilassiamo l'elemento con l'errore maggiore, questo metodo è più veloce: bastano 17 iterazioni per raggiungere l'equilibrio a meno di tre cifre decimali esatte. Non riportiamo i risultati per brevità: essi si possono trovare grazie al seguente notebook di Mathematica.

```
t=0.1
P={ {0,t/2,t/2,t/2,t/2,t/2}, {0,0,t,t,t,t}, {0,0,0,t,t,t}, {0,0,0,0,t,t}, {0,0,0,0,0,t}, {0,0,0,0,0,0} }
NN={ {1,0,0,0,0,0}, {-t,1,0,0,0,0}, {-t,-t,1,0,0,0}, {-t,-t,-t,1,0,0}, {-t,-t,-t,-t,1,0}, {-t,-t,-t,-t,-t,1} }
M=NN-P
e={0,0,0,0,0,1}
Ninv=Inverse[NN]
b={0,0,0,0,0,0}
maxr=1.
While[(maxr>0.001),
  r=N[e-M.b];
  maxr=Max[r[[1]], r[[2]], r[[3]], r[[4]], r[[5]],r[[6]]];
```

```

For [i=1, i<7, i++,
  If [(r[[i]]==maxr),i0=i,Continue]
];
b[[i0]]=N[(Ninv.P)[[i0]].b + Ninv[[i0]].e];
Print[N[b]] ]

```

Ecco le approssimazioni successive secondo Southwell ottenute tramite questo notebook di Mathematica:

```

{0.,0.,0.,0.,0.,1.}
{0.,0.,0.,0.,0.139755,1.}
{0.,0.,0.,0.144806,0.139755,1.}
{0.,0.,0.148367,0.144806,0.139755,1.}
{0.,0.150457,0.148367,0.144806,0.139755,1.}
{0.0791693,0.150457,0.148367,0.144806,0.139755,1.}
{0.0791693,0.150457,0.148367,0.144806,0.139755,1.06874}
{0.0791693,0.150457,0.148367,0.144806,0.163011,1.06874}
{0.0791693,0.150457,0.148367,0.163851,0.163011,1.06874}
{0.0791693,0.150457,0.16432,0.163851,0.163011,1.06874}
{0.0791693,0.164545,0.16432,0.163851,0.163011,1.06874}
{0.0862235,0.164545,0.16432,0.163851,0.163011,1.06874}
{0.0862235,0.164545,0.16432,0.163851,0.163011,1.07451}
{0.0862235,0.164545,0.16432,0.163851,0.165701,1.07451}
{0.0862235,0.164545,0.16432,0.165798,0.165701,1.07451}
{0.0862235,0.164545,0.165846,0.165798,0.165701,1.07451}
{0.0862235,0.165868,0.165846,0.165798,0.165701,1.07451}
{0.0868861,0.165868,0.165846,0.165798,0.165701,1.07451}

```

ESEMPIO 3: DISTRIBUZIONE DELLA LUCE ENTRO UNA STANZA RETTANGOLARE

Ora, invece di una stanza a forma di cubo, consideriamo una stanza rettangolare, di altezza 1, lunghezza L e profondità W . Le condizioni di illuminazione sono le stesse della stanza cubica vista precedentemente nell' [Esempio 2](#): le trascriviamo qui per completezza. La faccia alta ha una lampada che emette luce con intensità 1. Le altre facce non emettono luce propria, solo riflessa. Non ci sono finestre. La base, cioè il pavimento, è più scura degli altri muri: fissiamo $\rho_1 = 1/4$. Le pareti ed il soffitto sono grigio chiaro, e quindi poniamo $\rho_i = 1/2$ per $i=2, 3, 4, 5, 6$. Anche in questo caso ci limitiamo all'approssimazione rudimentale nella quale ciascuna parete forma un unico elemento (quindi non noteremo gradazioni di luce sulle pareti). Per comodità scriviamo ρ_+ invece di $1/2$ e ρ_- invece di $1/4$. Quindi il vettore dell'energia propria emessa è $\mathbf{e} = (0,0,0,0,0,1)$, mentre il vettore della riflettività è $\mathbf{p} = (1/4, 1/2, 1/2, 1/2, 1/2, 1/2)$. Nel seguito scriveremo \mathbf{e} invece di \mathbf{e} .

Anche in questo caso gli elementi sono piani, e quindi i fattori di forma diagonali, F_{i-i} , sono tutti nulli. Ora però la stanza non è completamente simmetrica. Ogni faccia copre lo stesso angolo solido solo quando è vista dalle due facce laterali opposte, cioè antistanti fra loro. Quindi ogni faccia ha un fattore di forma per la faccia antistante ed altri due per le due coppie di facce laterali. In tutto quindi ci sono 9 fattori di forma diversi. Ma la relazione di reciprocità della reversibilità dei percorsi ottici, introdotta nella sezione [L'equazione della radiosità](#), ci permette di ricavare F_{ji} da F_{ij} : $F_{ij} = A_j/A_i F_{ji}$. Perciò dobbiamo solo calcolare, tramite integrazione analitica o numerica, tre fattori di forma per facce in disposizione antistante e tre per disposizione laterale. Numeriamo le facce così: 1 il pavimento, 6 il soffitto, 2 e 4 le pareti di lunghezza L , 3 e 5 le pareti di lunghezza W . In realtà, l'equazione della conservazione dell'energia (anch'essa introdotta nella sezione [L'equazione della radiosità](#)) $\sum_j F_{ij} = 1$ ci fornisce il valore dell'ultimo fattore di forma F_{i-m} di ogni faccia i , una volta noto F_{ij} per tutti i j diversi da m . Quindi dobbiamo calcolare in tutto sei fattori di forma.

Calcoliamo i fattori dalla seguente formula, dimostrata nella sezione [Calcolo dei fattori di forma](#):

$$F_{i-j} = \frac{1}{A_i} \int_{A_i} \int_{A_j} \frac{\cos \theta_i \cos \theta_j}{\pi r^2} H_{ij} dA_j dA_i$$

Cominciamo con le facce opposte. Di nuovo, non ci sono ostruzioni dentro la stanza, quindi H vale costantemente 1, e le facce antistanti sono parallele, quindi $\theta_i = \theta_j$. Il calcolo è analogo a quello per la stanza cubica dell'[Esempio 2](#). Approssimiamo l'integrale numericamente, usando un notebook di [Mathematica](#) analogo a quello impiegato per la stanza cubica. Ad esempio, se $L=2$ e $W=2.5$, troviamo i seguenti fattori di forma per le facce antistanti, a meno di due cifre decimali:

$$F_{1-6} = F_{6-1} = 0.45$$

$$F_{2-4} = F_{4-2} = 0.13$$

$$F_{3-5} = F_{5-3} = 0.08$$

Invece, se $L=W=2$ (stanza quadrata) si trova:

$$F_{1-6} = F_{6-1} = 0.42$$

$$F_{2-4} = F_{4-2} = F_{3-5} = F_{5-3} = 0.12$$

Ecco il notebook che abbiamo usato per il caso $i=1, j=6$ (gli altri si ottengono ruotando in ordine ciclico i valori di height, length e width):

```
height=1
length=2
width =2

pointdown[x,y]={x,y,0};
pointup[x',y']={x',y',height};
displacement[x,y,x',y']=pointup[x',y']-pointdown[x,y];
NormSquare[x,y,x',y']=0;
For[i=1,i<4,i++,
  NormSquare[x,y,x',y']+=displacement[x,y,x',y'][[i]]^2];
Integrand[x,y,x',y']=height^2/ NormSquare[x,y,x',y']^2;
F=(1/(Pi * length * width))
NIntegrate[Integrand[x,y,x',y'],{x,0,length},{y,0,width},{x',0,length},
{y',0,width}]; Print[F]
```

Ora consideriamo il problema del calcolo di fattori di forma sulle facce laterali. Possiamo ancora supporre che la stanza abbia le facce parallele ai piani coordinati e giaccia nell'ottante positivo, con un vertice nell'origine. Allora la sua faccia 0 (il pavimento) ha coordinate $(x,y,0)$ per x e y che variano fra 0 e L e fra 0 e W , rispettivamente. La faccia laterale per $x=0$ ha coordinate $(0, y', z')$ per y' che varia fra 0 e L e z' fra 0 e 1. Se p è il vettore da $(x,y,0)$ a $(0, y', z')$ e q il versore $p/||p||$, allora nell'integrale $\cos\theta_1$ è il prodotto scalare fra q ed il versore $(0,0,1)$ normale al piano di base, mentre $\cos\theta_2$ è il prodotto scalare fra q ed il versore $(1,0,0)$ normale al piano $x=0$ che contiene l'elemento 2. Quindi $\cos\theta_1 = z'/||p||$ e $\cos\theta_2 = x/||p||$. Inoltre $r=||p||$. (Questi risultati sono stati già osservati nella sezione [Calcolo dei fattori di forma](#)). L'integrale si può calcolare analiticamente, ma il calcolo è laborioso. Oppure lo si può approssimare numericamente con Mathematica; per $L=2$ e $W=2.5$ si trova, ad esempio:

$$F_{1-2} = 0.23 = A_2/A_1 F_{2-1}$$

Però l'approssimazione numerica è affetta da un forte errore, perché l'integrando diverge quando r si avvicina a 0, cioè al segmento in cui il pavimento tocca la parete. Per evitare rischi di imprecisione, supponiamo che la stanza sia quadrata, cioè che sia $L=W$ (ma il valore comune non sia 1, perché abbiamo già trattato la stanza cubica). Allora $F_{1-2}=F_{1-3}=F_{1-4}=F_{1-5}=F_{6-2}=F_{6-3}=F_{6-4}=F_{6-5}$, e dall'equazione della conservazione dell'energia (o

dell'angolo solido totale) $\sum_j F_{ij} = 1$ (introdotta nella sezione [L'equazione della radiosità](#)) si ricava il loro valore comune, che chiameremo $F_{\text{pav-lato}}$:

$$F_{\text{pav-lato}} = (1 - F_{1-6})/4 = 0.145$$

Analogamente, chiamiamo $F_{\text{lato-pav}}$ il valore comune $F_{2-1} = F_{2-6} = F_{4-1} = F_{4-6} = F_{3-1} = F_{3-6} = F_{5-1} = F_{5-6}$ e $F_{\text{lato-lato}}$ il valore comune $F_{3-2} = F_{5-2} = F_{3-4} = F_{5-4} = F_{2-3} = F_{2-5} = F_{4-3} = F_{4-5}$. Sappiamo che $F_{2-1} = A_1/A_2 F_{1-2} = 2 F_{1-2}$, quindi $F_{\text{lato-pav}} = 2 F_{\text{pav-lato}} = 0.29$. D'altra parte, sappiamo che $\sum_j F_{2-j} = 1$, ma $F_{2-3} = F_{2-5} = F_{\text{lato-lato}}$, $F_{2-4} = 0.12$ e $F_{2-6} = F_{2-1} = F_{\text{lato-pav}} = 0.29$. Da qui segue che $F_{\text{lato-lato}} = (1 - 0.12 - 2 \times 0.29)/2 = 0.15$. Abbiamo determinato tutti i fattori di forma per la stanza quadrata. Riassumendo:

ci sono 2 fattori di forma frontali:

da pavimento a soffitto (o viceversa)	0.42
da parete a parete antistante	0.12
c' è un fattore di forma da parete a parete adiacente	0.15
c' è un fattore di forma da parete a pavimento (o soffitto)	0.29
c' è un fattore di forma da pavimento (o soffitto) a parete	0.145

Con questi dati possiamo scrivere la matrice M del sistema della radiosità, e le matrici triangolare inferiore e superiore N e P della decomposizione $M = N - P$, come in (6') nella sezione [Analisi del metodo di Gauss-Seidel](#).

Ci limitiamo a scrivere N e P , ricordando che la riflettività del pavimento vale $\frac{1}{4}$ e quella degli altri elementi $\frac{1}{2}$.

$$N = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ -\frac{1}{2}0.29 & 1 & 0 & 0 & 0 & 0 \\ -\frac{1}{2}0.29 & -\frac{1}{2}0.15 & 1 & 0 & 0 & 0 \\ -\frac{1}{2}0.29 & -\frac{1}{2}0.12 & -\frac{1}{2}0.15 & 1 & 0 & 0 \\ -\frac{1}{2}0.29 & -\frac{1}{2}0.15 & -\frac{1}{2}0.12 & -\frac{1}{2}0.15 & 1 & 0 \\ -\frac{1}{2}0.42 & -\frac{1}{2}0.145 & -\frac{1}{2}0.145 & -\frac{1}{2}0.145 & -\frac{1}{2}0.145 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ -0.145 & 1 & 0 & 0 & 0 & 0 \\ -0.145 & -0.075 & 1 & 0 & 0 & 0 \\ -0.145 & -0.06 & -0.075 & 1 & 0 & 0 \\ -0.145 & -0.075 & -0.06 & -0.075 & 1 & 0 \\ -0.21 & -0.0725 & -0.0725 & -0.0725 & -0.0725 & 1 \end{pmatrix}$$

$$P = \begin{pmatrix} 0 & \frac{1}{2}0.145 & \frac{1}{2}0.145 & \frac{1}{2}0.145 & \frac{1}{2}0.145 & \frac{1}{4}0.42 \\ 0 & 0 & \frac{1}{2}0.15 & \frac{1}{2}0.15 & \frac{1}{2}0.15 & \frac{1}{4}0.29 \\ 0 & 0 & 0 & \frac{1}{2}0.15 & \frac{1}{2}0.15 & \frac{1}{4}0.29 \\ 0 & 0 & 0 & 0 & \frac{1}{2}0.15 & \frac{1}{4}0.29 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{4}0.29 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} 0 & 0.0725 & 0.0725 & 0.0725 & 0.0725 & 0.1005 \\ 0 & 0 & 0.075 & 0.06 & 0.075 & 0.0725 \\ 0 & 0 & 0 & 0.075 & 0.06 & 0.0725 \\ 0 & 0 & 0 & 0 & 0.075 & 0.0725 \\ 0 & 0 & 0 & 0 & 0 & 0.0725 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

1. Metodo di Jacobi

Ora possiamo eseguire le iterazioni come nei precedenti esempi della [tenda piramidale](#) e della [stanza cubica](#). Per non diventare ripetitivi, ci limitiamo a riportare i risultati del metodo di Jacobi (questa volta alle prime quattro cifre decimali), lasciando al lettore i calcoli per questo metodo (per il quale però trascriviamo il notebook di [Mathematica](#)) e per il metodo di Gauss-Seidel. Il metodo di Southwell verrà implicitamente trattato fra qualche pagina quando illustreremo il raffinamento progressivo.

$$b^{(1)} = (0.1005, 0.0871, 0.0936, 0.0993, 0.1067, 1.0491)$$

$$b^{(2)} = (0.1335, 0.1164, 0.118, 0.1192, 0.1202, 1.0624)$$

$$b^{(3)} = (0.1411, 0.1225, 0.1228, 0.1231, 0.1233, 1.0653)$$

$$b^{(4)} = (0.1427, 0.1238, 0.1238, 0.1239, 0.1239, 1.0659)$$

$$b^{(5)} = (0.1430, 0.124, 0.124, 0.1241, 0.1241, 1.066)$$

$$b^{(6)} = (0.1431, 0.1241, 0.1241, 0.1241, 0.1241, 1.066)$$

$$b^{(7)} = (0.1431, 0.1241, 0.1241, 0.1241, 0.1241, 1.066)$$

Ecco il notebook che abbiamo usato per il calcolo:

```
t1=0.0725; t2=0.075; t3=0.06; t4=0.145
P={ {0,t1,t1,t1,t1,0.1005}, {0,0,t2,t3,t2,t1}, {0,0,0,t2,t3,t1}, {0,0,0,0,t2,t1}, {0,0,0,0,0,t1}, {0,0,0,0,0,0} }
NN={ {1,0,0,0,0,0}, {-t4,1,0,0,0,0}, {-t4,-t2,1,0,0,0}, {-t4,-t3,-t2,1,0,0}, {-t4,-t2,-t3,-t2,1,0}, {-0.21,-t1,-t1,-t1,-t1,1} }; TraditionalForm[NN]
e={0,0,0,0,0,1}
Ninv=Inverse[NN]; TraditionalForm[Ninv]
b={0,0,0,0,0,0}
For[i=0, i<10, i++, N[Expand[b=Ninv .P. b + Ninv .e]]; Print[b] ]
```

Si noti quanto rapido sia l'aumento delle componenti del vettore di radiosità nelle prime iterazioni: infatti all'inizio l'iterazione cambia la prima cifra decimale. Poi la prima cifra si stabilizza e comincia a cambiare la seconda, e via via le altre. Quindi le differenze successive diventano via via ordini di grandezza più piccoli. In effetti, osserviamo che le prime due cifre decimali sono stabili già dopo tre iterazioni, le prime tre dopo quattro iterazioni, le prime quattro dopo sei iterazioni.

Proprio per questo rapido incremento iniziale e lento incremento successivo il metodo del raffinamento progressivo, introdotto nelle sezioni [Verso un raffinamento progressivo](#) e seguenti, permette di estrapolare in pochissime iterazioni il punto fisso finale. Tra qualche pagina svolgeremo i calcoli con questo metodo.

2. Metodo di Southwell

In questo metodo, utilizziamo il vettore dell'energia creata $\underline{\varepsilon}$, quello dell'energia irradiata $\underline{\beta}$, e la matrice dell'energia \underline{K} . Le pareti laterali hanno area 2, mentre pavimento ($i=1$) e

soffitto ($i=6$) hanno area 4. Poiché $K_{ij} = (A_i/A_j)M_{ij}$ la matrice K si scompone come $K = N' - P'$, dove le matrici N' e P' differiscono da N e P precedentemente utilizzate in questo modo:

- N' differisce da N solo nella ultima riga, colonne 2,3,4,5 (dove i coefficienti vengono moltiplicati per $1/2$), e nella prima colonna, righe 2,3,4,5 (i cui coefficienti vengono moltiplicati per 2)
- P' differisce da P solo nella prima riga, colonne 2,3,4,5 (dove i coefficienti vengono moltiplicati per 2), e nella ultima colonna, righe 2,3,4,5 (i cui coefficienti vengono moltiplicati per $1/2$)

Ora scriviamo per esteso le nuove matrici. Con queste matrici ed il notebook di Mathematica usato prima si possono trovare gli approssimanti iterativi $\beta^{(k)}$, e da essi ricavare $b^{(k)}$. Per il momento omettiamo i calcoli, perché li ripeteremo fra qualche pagina nell'illustrare il raffinamento progressivo.

$$\begin{aligned}
 N' &= \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ -0.29 & 1 & 0 & 0 & 0 & 0 \\ -0.29 & -(1/2)0.15 & 1 & 0 & 0 & 0 \\ -0.29 & -(1/2)0.12 & -(1/2)0.15 & 1 & 0 & 0 \\ -0.29 & -(1/2)0.15 & -(1/2)0.12 & -(1/2)0.15 & 1 & 0 \\ -(1/2)0.42 & -(1/4)0.145 & -(1/4)0.145 & -(1/4)0.145 & -(1/4)0.145 & 1 \end{pmatrix} \\
 &= \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ -0.29 & 1 & 0 & 0 & 0 & 0 \\ -0.29 & -0.075 & 1 & 0 & 0 & 0 \\ -0.29 & -0.06 & -0.075 & 1 & 0 & 0 \\ -0.29 & -0.075 & -0.06 & -0.075 & 1 & 0 \\ -0.21 & -0.03625 & -0.03625 & -0.03625 & -0.03625 & 1 \end{pmatrix} \\
 P' &= \begin{pmatrix} 0 & 0.145 & 0.145 & 0.145 & 0.145 & (1/4)0.42 \\ 0 & 0 & (1/2)0.15 & (1/2)0.12 & (1/2)0.15 & (1/8)0.29 \\ 0 & 0 & 0 & (1/2)0.15 & (1/2)0.12 & (1/8)0.29 \\ 0 & 0 & 0 & 0 & (1/2)0.15 & (1/8)0.29 \\ 0 & 0 & 0 & 0 & 0 & (1/8)0.29 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \\
 &= \begin{pmatrix} 0 & 0.145 & 0.145 & 0.145 & 0.145 & 0.1005 \\ 0 & 0 & 0.075 & 0.06 & 0.075 & 0.03625 \\ 0 & 0 & 0 & 0.075 & 0.06 & 0.03625 \\ 0 & 0 & 0 & 0 & 0.075 & 0.03625 \\ 0 & 0 & 0 & 0 & 0 & 0.03625 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}
 \end{aligned}$$

3. Raffinamento progressivo

Come già spiegato nella sottosezione *Interpretazione fisica del procedimento di Southwell: luce emessa nell'ambiente*, un vantaggio del metodo di che stiamo presentando (Southwell con raffinamento) è che si presta a riconsiderare il sistema lineare della radiosità dal punto di vista nel quale si considera l'energia irradiata da ogni elemento nell'ambiente invece che ricevuta dall'ambiente. In questo modo si ottengono cospicui guadagni di tempo e di allocazione di memoria nel calcolo dei fattori di forma, come visto nella sottosezione *Vantaggi numerici dell'approccio ad emissione di energia nell'ambiente*. Poiché però i fattori di forma li abbiamo già calcolati, tralasciamo questo aspetto e ci rivolgiamo alla estrapolazione del punto di equilibrio tramite raffinamento progressivo degli approssimanti. Questo richiede una stima della energia residua da irradiare, pesata con le aree degli elementi, presentata nella sottosezione *Correzione ambientale della luminosità nel corso delle iterazioni*, alla quale facciamo riferimento per le formule che stiamo per citare.

Rammentiamo gli ingredienti di questa stima. Gli approssimanti della radiosità vengono incrementati di un termine che compensa l'energia ancora da irradiare in seguito ad ulteriori riflessioni, dato da (37):

$$b_i^{(k)'} = b_i^{(k)} + \rho_i \Phi$$

dove Φ è una stima del flusso per unità di area dell'energia totale non irradiata, dato da (36):

$$\Phi = RH/A.$$

In quest'ultima equazione, A è la somma totale delle aree degli elementi (nel nostro caso $A=16$, perché pavimento e soffitto hanno area 4 e le pareti hanno area 2), R è la riflettività media ambientale che tiene conto di un numero arbitrario di interriflessioni della luce, e che si ottiene a partire dalla media pesata della riflettività degli elementi, $\rho_{media} = \sum_{i=1, \dots, n} \rho_i A_i / \sum_{i=1, \dots, N} A_i$, grazie a (34):

$$R = 1/(1 - \rho_{media}).$$

Infine, $H = H^{(k)}$ è una stima dell'energia totale ancora non irradiata nell'ambiente dopo

$$H^{(k)} = \sum_{i=1, \dots, n} r_i^{(k)}$$

con il termine residuo dell'energia dato da (30):

$$r_i^{(k)} = \beta_i^{(k+1)} - \beta_i^{(k)}.$$

che è equivalente a $r^{(k)} = \varepsilon - K \beta^{(k)}$

Nel nostro esempio, il pavimento ha riflettività $\frac{1}{4}$ ed area 4, le altre pareti ed il soffitto hanno riflettività $\frac{1}{2}$ ed area rispettivamente 2 e 4, quindi $\rho_{media} = (4 \times \frac{1}{4} + (4 \times 2 + 4) \times \frac{1}{2})/16 = 7/16 = 0.4375$. Pertanto $R = 16/9 = 1.778$, e diventa facile trovare $b_i^{(k)'}$ da (37):

basta generare le iterazioni $\beta_i^{(k)}$ col metodo di Southwell, calcolare $r_i^{(k)}$ da (30), quindi $H^{(k)}$ da (35) ed infine $\Phi = \Phi^{(k)}$ da (36). Tutto questo si ottiene quindi con questa semplice modifica del notebook con cui abbiamo generato le iterazioni del metodo di Southwell:

```
A={4,2,2,2,2,4};
rho={1/4,1/2,1/2,1/2,1/2,1/2};t=0.03625;t1=0.075;t2=0.06;t3=0.21;t4=0.29;
t5=0.1005;
P={0,t4/2,t4/2,t4/2,t4/2,t5},{0,0,t1,t2,t1,t
},{0,0,0,t1,t2,t},{0,0,0,0,t1,t},{0,0,0,0,0,t},{0,0,0,0,0,0}};
NN={1,0,0,0,0,0},{-t4,1,0,0,0,0},{-t4,-t1,1,0,0,0},{-t4,-t2,-t1,1,0,0},{-t4,-
t1,-t2,-t1,1,0},
{-t3,-t,-t,-t,-t,1}};
K=NN-P;
e={0,0,0,0,0,1};
Ninv=Inverse[NN]; TraditionalForm[NN];
b={0,0,0,0,0,0};
eps=e; beta=b;bprime=b;
AreaTot=0;
For [i=1,i<7,i++,AreaTot=AreaTot+A[[i]]];
rhoMedia=0;
For [i=1,i<7,i++,rhoMedia=rhoMedia+rho[[i]] A[[i]] ];
rhoMedia=rhoMedia/AreaTot;
R=1/(1-rhoMedia);
kk=0;maxr=1;
While[(maxr>0.001),
  r=N[eps-K.beta];
  maxr=Max[r[[1]], r[[2]], r[[3]], r[[4]], r[[5]],r[[6]]];
  For [i=1, i<7, i++,
    If [(r[[i]]==maxr),i0=i,Continue]
  ];
  beta[[i0]]=N[(Ninv .P)[[i0]]. beta+ Ninv[[i0]] .eps];
  For [i=1,i<7,i++,b[[i]]=beta[[i]]/A[[i]]];
  H=0;
  For [i=1,i<7,i++,H=H+r[[i]]];
  Phi=R H/AreaTot;
  For [i=1,i<7,i++,bprime[[i]]=b[[i]]+ rho[[i]] Phi];
  Print[kk]; Print[N[b]];Print[N[bprime]];
  kk++]
```

Alla trentatreesima iterazione i due metodi danno lo stesso risultato, e da quel momento in poi le prime quattro cifre decimali si stabilizzano. Il metodo numerico basato sul raffinamento è ad ogni iterazione considerevolmente più vicino al risultato finale di quanto non sia il metodo di Southwell. Nonostante questo, entrambi i metodi impiegano lo stesso numero di iterazioni a stabilizzarsi, ma già alla sesta iterazione il metodo del raffinamento fornisce una buona approssimazione del risultato finale. A differenza del metodo di Southwell, il raffinamento non produce una approssimazione monotona crescente, anzi, dopo una crescita notevolissima al primo passo, gli approssimanti del raffinamento convergono in maniera decrescente, ma ad ogni passo essi sono più grandi di quelli di Southwell, come deve essere a causa di (37).

Ecco le iterazioni successive del metodo di Southwell ricavate da Mathematica. Per ogni iterazione, il primo vettore è l'approssimante vero, il secondo l'approssimante aumentato nel senso del raffinamento progressivo:

```
1
{0.,0.,0.,0.,0.,0.25}
{0.0277778,0.0555556,0.0555556,0.0555556,0.0555556,0.305556}
2
{0.025125,0.,0.,0.,0.,0.25}
```

{0.0319444,0.0136389,0.0136389,0.0136389,0.0136389,0.263639}
3
{0.025125,0.,0.,0.,0.040056,0.25}
{0.0329774,0.0157047,0.0157047,0.0157047,0.0557607,0.265705}
4
{0.025125,0.,0.,0.0425984,0.040056,0.25}
{0.0316227,0.0129954,0.0129954,0.0555938,0.0530514,0.262995}
5
{0.025125,0.,0.0449013,0.0425984,0.040056,0.25}
{0.030182,0.0101141,0.0550154,0.0527125,0.05017,0.260114}
6
{0.025125,0.0469889,0.0449013,0.0425984,0.040056,0.25}
{0.0286635,0.0540659,0.0519783,0.0496754,0.047133,0.257077}
7
{0.0377795,0.0469889,0.0449013,0.0425984,0.040056,0.25}
{0.0397288,0.0508876,0.0488001,0.0464971,0.0439547,0.253899}
8
{0.0377795,0.0469889,0.0449013,0.0425984,0.040056,0.261529}
{0.0402491,0.0519281,0.0498405,0.0475376,0.0449952,0.266468}
9
{0.0377795,0.0469889,0.0449013,0.0425984,0.0522409,0.261529}
{0.0392826,0.0499951,0.0479076,0.0456046,0.0552471,0.264535}
10
{0.0377795,0.0469889,0.0449013,0.0530102,0.0522409,0.261529}
{0.0388705,0.049171,0.0470834,0.0551923,0.0544229,0.263711}
11
{0.0377795,0.0469889,0.0535775,0.0530102,0.0522409,0.261529}
{0.0385184,0.0484667,0.0550553,0.054488,0.0537187,0.263006}
12
{0.0412054,0.0469889,0.0535775,0.0530102,0.0522409,0.261529}
{0.0416509,0.0478799,0.0544684,0.0539012,0.0531318,0.26242}
13
{0.0412054,0.053977,0.0535775,0.0530102,0.0522409,0.261529}
{0.0417917,0.0551496,0.0547501,0.0541829,0.0534135,0.262701}
14
{0.0412054,0.053977,0.0535775,0.0530102,0.0522409,0.262703}
{0.0415554,0.0546769,0.0542774,0.0537102,0.0529409,0.263403}
15
{0.0412054,0.053977,0.0535775,0.0530102,0.0547526,0.262703}
{0.0414569,0.05448,0.0540805,0.0535133,0.0552557,0.263206}
16
{0.0420122,0.053977,0.0535775,0.0530102,0.0547526,0.262703}
{0.0421788,0.0543101,0.0539107,0.0533434,0.0550858,0.263036}
17
{0.0420122,0.053977,0.0535775,0.0549109,0.0547526,0.262703}
{0.0422119,0.0543765,0.053977,0.0553104,0.0551521,0.263103}
18
{0.0420122,0.053977,0.0550148,0.0549109,0.0547526,0.262703}
{0.0421476,0.0542479,0.0552858,0.0551819,0.0550235,0.262974}
19
{0.0420122,0.0550806,0.0550148,0.0549109,0.0547526,0.262703}
{0.042099,0.0552543,0.0551886,0.0550846,0.0549263,0.262877}
20
{0.0423342,0.0550806,0.0550148,0.0549109,0.0547526,0.262703}
{0.0423837,0.0551797,0.0551139,0.05501,0.0548517,0.262802}
21
{0.0423342,0.0550806,0.0550148,0.0549109,0.0551788,0.262703}
{0.042397,0.0552062,0.0551404,0.0550365,0.0553043,0.262829}
22
{0.0423342,0.0550806,0.0550148,0.0549109,0.0551788,0.262898}
{0.0423825,0.0551773,0.0551116,0.0550076,0.0552755,0.262995}
23
{0.0423342,0.0550806,0.0550148,0.0552347,0.0551788,0.262898}

```

{0.0423662,0.0551446,0.0550789,0.0552988,0.0552428,0.262962}
24
{0.0423342,0.0550806,0.0552528,0.0552347,0.0551788,0.262898}
{0.0423552,0.0551227,0.055295,0.0552768,0.0552209,0.26294}
25
{0.0424254,0.0550806,0.0552528,0.0552347,0.0551788,0.262898}
{0.0424384,0.0551066,0.0552789,0.0552607,0.0552048,0.262924}
26
{0.0424254,0.0552633,0.0552528,0.0552347,0.0551788,0.262898}
{0.0424422,0.0552968,0.0552864,0.0552683,0.0552123,0.262932}
27
{0.0424254,0.0552633,0.0552528,0.0552347,0.0552819,0.262898}
{0.042436,0.0552845,0.055274,0.0552559,0.055303,0.262919}
28
{0.0424461,0.0552633,0.0552528,0.0552347,0.0552819,0.262898}
{0.0424532,0.0552775,0.055267,0.0552489,0.0552961,0.262912}
29
{0.0424461,0.0552633,0.0552528,0.0552884,0.0552819,0.262898}
{0.0424541,0.0552792,0.0552687,0.0553043,0.0552978,0.262914}
30
{0.0424461,0.0552633,0.0552528,0.0552884,0.0552819,0.262923}
{0.0424523,0.0552756,0.0552651,0.0553006,0.0552941,0.262935}
31
{0.0424461,0.0552633,0.0552948,0.0552884,0.0552819,0.262923}
{0.0424502,0.0552714,0.0553029,0.0552965,0.05529,0.262931}
32
{0.0424461,0.0552967,0.0552948,0.0552884,0.0552819,0.262923}
{0.0424488,0.055302,0.0553001,0.0552936,0.0552871,0.262928}
33
{0.042458,0.0552967,0.0552948,0.0552884,0.0552819,0.262923}
{0.0424595,0.0552997,0.0552978,0.0552914,0.0552849,0.262926}

```

Queste approssimazioni nelle quali ciascuna faccia e' un elemento sono terribilmente rudimentali. Per una ragionevole gradazione delle luci, una faccia dovrebbe scomporsi in molte centinaia di elementi. Ma la mole di calcoli crescerebbe notevolmente. Supponiamo ad esempio di considerare la modellazione, ancora troppo rudimentale, in cui ciascuna faccia si scompone in 25 elementi (quindi l'elemento centrale del soffitto è la lampada: finalmente il soffitto non emette più luce uniformemente da tutta la propria superficie!). Ci sono in tutto $25 \times 6 = 150$ elementi. Quindi le matrici sono di dimensione 150. Quanti fattori di forma devono essere calcolati? A prima vista sembrerebbe di doverne calcolare $150^2 = 22500$, un numero già elevatissimo. Ma invece sono di meno, perché molti coincidono. Ad esempio, data un elemento, tutti i fattori di forma verso elementi di un'altra parete dipendono solo dallo spostamento relativo fra le due, quindi sono solo 25; inoltre, l'insieme dei fattori di forma da elementi di una parete a elementi di una parete adiacente rimane identico indipendentemente da quale delle quattro pareti adiacenti si sceglie. Quindi basta calcolare 25 fattori di forma per elementi su pareti adiacenti ed altri 25 per elementi su pareti antistanti: in tutto 50. Però poi bisogna scrivere il codice in maniera intelligentemente appropriata per fare in modo che tenga conto in modo automatico di quali fattori di forma si ripetono.

Lasciamo al lettore il calcolo di quale mole di lavoro sarebbe richiesta se anziché suddividere ogni faccia in 25 elementi la considerassimo un unico elemento ma provvedessimo a sottostrutturarla in 25 sottoelementi.

ESERCIZI FINALI SULLA RADIOSITÀ

Esercizio 1. Una stanza di forma cubica ha una luce di intensità 1 sul soffitto ed una finestra che lascia entrare luce con intensità 1 sulla parete sud. Il fattore di forma fra pavimento e soffitto vale 0.2. Le pareti senza finestra ed il pavimento sono neri (coefficiente di riflettività zero); il soffitto e la parete sud hanno coefficiente di riflettività $\frac{1}{2}$. Calcolare le prime due iterazioni del metodo di Jacobi e del metodo di Gauss-Seidel per il calcolo delle radiosità, prendendo le pareti, il pavimento ed il soffitto come elementi.

Esercizio 2. Una scena consiste dell'interno di una semisfera di raggio 2, al centro della quale c'è una cabina semisferica di raggio 1. La scena è l'intercapedine fra le due semisfere. La luce viene da un'apertura sulla semisfera esterna, con potenza 1. La semisfera esterna ha fattore di riflettività $\frac{2}{3}$. Quella interna ha fattore di riflettività $\frac{1}{3}$. Il pavimento ha fattore di riflettività 0. Il fattore di forma dal pavimento alla semisfera esterna è $\frac{7}{8}$.

- Trovare tutti gli altri fattori di forma.
- Trovare le prime tre iterazioni della radiosità ($b^{(1)}$, $b^{(2)}$ e $b^{(3)}$) con il metodo di Jacobi.
- Trovare $b^{(1)}$ con il metodo di Gauss-Seidel.
- Trovare la seconda iterazione del metodo di Southwell.

Esercizio 3. La scena consiste dell'interno dell'astronave Enterprise, fatta a ciambella, di metallo con fattore di riflettività $\frac{1}{2}$. Tutte le luci sono spente a causa di un attacco dei Klingon, tranne che nel pannello di comando, che emette luce con potenza 1. I membri dell'equipaggio, impavidi nonostante l'attacco, rimasti senza computers funzionanti, alla fioca luce emessa dal pannello di comando sono impegnati a calcolare a mano l'illuminazione che si distribuisce dal pannello di comando all'astronave (un modo come un altro per scoprire se nella corazza si è formato qualche buco, che ovviamente avrebbe riflettività 0 invece che $\frac{1}{2}$). Essi immaginano l'interno dell'astronave separato in quattro settori di 90 gradi tramite piani perpendicolari che si incontrano sull'asse centrale della ciambella (quello rispetto al quale la ciambella è invariante per rotazione). Così l'astronave si scompone in quattro elementi, al centro di una delle quali è il pannello di controllo. La ciambella è sufficientemente sottile perché da ciascun elemento si riescano a vedere solo quelle adiacenti. Il fattore di forma di ogni elemento su sé stesso è $\frac{1}{2}$. Purtroppo, quando comincia il calcolo l'intero elemento opposto al pannello di comando è stata demolito dai Klingon, e quindi è diventato nero (il colore dello spazio esterno). I membri dell'equipaggio, prima di essere uccisi dai Klingon, sono in grado di:

- Trovare tutti i fattori di forma.
- Trovare le prime tre iterazioni della radiosità ($b^{(1)}$, $b^{(2)}$ e $b^{(3)}$) con il metodo di Jacobi.
- Trovare $b^{(1)}$ con il metodo di Gauss-Seidel.

- Trovare la seconda iterazione del metodo di Southwell.

Lo siete anche voi, prima di essere uccisi da questo esercizio?

Assumete di sapere che l'elemento opposto al pannello di comando sia ormai nero.

Esercizio 4. Una scena consiste dell'interno di una torre cilindrica di raggio r ed altezza m . Sia α il fattore di forma dal pavimento alla parete cilindrica. È possibile determinare tutti i fattori di forma in funzione di α ? Se non lo è, mostrare perché; se lo è, per quali r e m la scelta $\alpha = 1/2$ è compatibile con il problema? Se lo è per $r = m$ (m come definito sopra), scrivere in tal caso il sistema lineare della radiosità assumendo che il pavimento sia nero, la parete ed il soffitto bianchi (ossia con fattore di riflettività 1) ed il soffitto emetta luce a potenza 1, e calcolare le prime tre iterazioni per la radiosità con il metodo di Jacobi.

Svolgimento. Supponiamo, per maggiore generalità, che il raggio sia r (in seguito porremo $r=m$). Siano β il fattore di forma da pavimento a soffitto (e viceversa), γ quello da parete a pavimento (o soffitto), e σ quello dalla parete e sè stessa: non ci sono altri fattori di forma. L'area della parete è $2\pi rm$, quella del pavimento è πr^2 , quindi, per la regola di reciprocità,

$$\gamma = \frac{r}{2m} \alpha \quad (1)$$

Le identità relative alla conservazione dell'angolo solido totale sono $\alpha + \beta = 1$ e $2\gamma + \sigma = 1$. Grazie all'identità (1), queste ora possono essere riscritte come

$$\sigma = 1 - \frac{r}{m} \alpha \quad (2)$$

$$\beta = 1 - \alpha \quad (3)$$

e quindi tutti i fattori di forma sono funzioni di α . Affinché il risultato sia accettabile occorre che tutti i fattori di forma siano positivi e non superiori a 1. Quindi, per (2), la scelta $\alpha = 1/2$ è compatibile se e solo se $r/m \leq 2$. In particolare lo è per $r = m$. In tal caso, numerando con 1 il pavimento, 2 la parete e 3 il soffitto, e ponendola matrice M della radiosità diventa

$$M = \begin{pmatrix} 1 & 0 & 0 \\ -\rho\gamma & 1-\rho\sigma & -\rho\gamma \\ -\rho\beta & -\rho\alpha & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ -\frac{\rho\alpha}{2} & 1-\rho(1-\alpha) & -\frac{\rho\alpha}{2} \\ -\rho(1-\alpha) & -\rho\alpha & 1 \end{pmatrix}$$

Ora spezziamo $M = N - P$, con

$$N = \begin{pmatrix} 1 & 0 & 0 \\ -\frac{\rho\alpha}{2} & 1-\rho(1-\alpha) & 0 \\ -\rho(1-\alpha) & -\rho\alpha & 1 \end{pmatrix} \quad \text{e} \quad P = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & \frac{\rho\alpha}{2} \\ 0 & 0 & 0 \end{pmatrix}$$

Qui N è del tipo $N = \begin{pmatrix} 1 & 0 & 0 \\ x & a & 0 \\ z & y & 1 \end{pmatrix}$, e si verifica subito che la sua inversa è

$$N^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ -\frac{x}{a} & \frac{1}{a} & 0 \\ \frac{xy-az}{a} & -\frac{y}{a} & 1 \end{pmatrix}$$

quindi nel nostro caso

$$N^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ \frac{\frac{1}{2}\rho^2\alpha^2 + a\rho(1-\alpha)}{1-\rho(1-\alpha)} & \frac{1}{1-\rho(1-\alpha)} & 0 \\ \frac{\rho\alpha}{1-\rho(1-\alpha)} & \frac{\rho\alpha}{1-\rho(1-\alpha)} & 1 \end{pmatrix} \quad N^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ \frac{1}{2} \frac{\rho\alpha}{1-\rho(1-\alpha)} & \frac{1}{1-\rho(1-\alpha)} & 0 \\ \frac{\rho(1-\alpha) + \frac{1}{2}\rho^2(1-2\alpha)}{1-\rho(1-\alpha)} & \frac{\rho\alpha}{1-\rho(1-\alpha)} & 1 \end{pmatrix}$$

Pertanto si ha

$$N^{-1}P = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} \frac{\rho\alpha}{1-\rho(1-\alpha)} \\ 0 & 0 & \frac{1}{2} \frac{\rho^2\alpha^2}{1-\rho(1-\alpha)} \end{pmatrix}$$

La potenza creata per unità di area è il vettore $\mathbf{e} = (0,0,1)$. Partendo dall'approssimazione iniziale $\mathbf{b}^0 = (0,0,0)$ si ottiene $\mathbf{b}^{(1)} = N^{-1}\mathbf{e} = (0,0,1)$ (la terza colonna della matrice N^{-1} !). Quindi $N^{-1}P \mathbf{b}^{(1)}$ è la terza colonna della matrice $N^{-1}P$, e pertanto

$$\mathbf{b}^{(2)} = N^{-1}P \mathbf{b}^{(1)} + N^{-1}\mathbf{e} = (N^{-1}P + \text{Id})\mathbf{e} = \frac{1}{2} \begin{pmatrix} 0 \\ \frac{\rho\alpha}{1-\rho(1-\alpha)} \\ \frac{\rho^2\alpha^2}{1-\rho(1-\alpha)} \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ \frac{1}{2} \frac{\rho\alpha}{1-\rho(1-\alpha)} \\ 1 + \frac{1}{2} \frac{\rho^2\alpha^2}{1-\rho(1-\alpha)} \end{pmatrix}$$

Allora

$$\begin{aligned}
b^{(3)} &= N^{-1} P b^{(2)} + N^{-1} e = N^{-1} P \begin{pmatrix} 0 \\ \frac{1}{2} \frac{\rho\alpha}{1-\rho(1-\alpha)} \\ 1 + \frac{1}{2} \frac{\rho^2\alpha^2}{1-\rho(1-\alpha)} \end{pmatrix} + e \\
&= \begin{pmatrix} 0 \\ \frac{1}{2} \frac{\rho\alpha}{1-\rho(1-\alpha)} + \frac{1}{4} \frac{\rho^3\alpha^3}{(1-\rho(1-\alpha))^2} \\ \frac{1}{2} \frac{\rho^2\alpha^2}{1-\rho(1-\alpha)} + \frac{1}{4} \frac{\rho^4\alpha^4}{(1-\rho(1-\alpha))^2} \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ \frac{1}{2} \frac{\rho\alpha}{1-\rho(1-\alpha)} + \frac{1}{4} \frac{\rho^3\alpha^3}{(1-\rho(1-\alpha))^2} \\ 1 + \frac{1}{2} \frac{\rho^2\alpha^2}{1-\rho(1-\alpha)} + \frac{1}{4} \frac{\rho^4\alpha^4}{(1-\rho(1-\alpha))^2} \end{pmatrix}
\end{aligned}$$

Nel nostro caso, $\rho = \alpha = 1/2$, e quindi

$$b^{(1)} = (0, 0, 1),$$

$$b^{(2)} = (0, 1/6, 25/24) = (0, 0.166667, 1.041667),$$

$$b^{(3)} = (0, 1/6 + 1/256, 1 + 1/24 + 1/1024) = (0, 0.170588, 1.042643).$$

Esercizio 5. Consideriamo una stanza cubica di lato 10 metri nella quale, per semplicità, assumiamo uguali tutti i fattori di forma non nulli. Nella stanza stendiamo un telo verticale, dal lato che congiunge due pareti a quello opposto (ossia quello che congiunge le due pareti opposte). Il telo divide la stanza in due parti uguali: consideriamo solo una delle due. Il telo e le pareti di questa metà della stanza sono elementi. Numeriamo il soffitto (triangolare) con 1, le due pareti (quadrato) con 2 e 3, il pavimento (triangolare) con 4, il telo con 5. Il telo è nero, le pareti ed il soffitto sono bianchi.

a. Supponiamo che il fattore di forma da soffitto a pavimento sia $1/4$, e lo stesso valga per il fattore di forma da parete a parete e da soffitto a parete. Questi dati sono sufficienti per calcolare tutti i fattori di forma? Sono compatibili fra loro? Se sì calcolarli, se no spiegare perché.

b. Come cambia la risposta se si chiede solo che il fattore di forma da soffitto a pavimento e quello da parete a parete siano $1/4$?

Svolgimento. Poiché gli elementies sono piane, i fattori di forma da un elemento a sé stessa sono nulli. Per simmetria, ci sono solo cinque fattori di forma indipendenti: da triangolo (soffitto) a quadrato (parete intera), da triangolo a triangolo (soffitto a pavimento), da rettangolo (telo) a triangolo, da quadrato a rettangolo e da quadrato a quadrato: i fattori di forma in senso inverso si ricavano da questi tramite l'equazione di reciprocità $A_i F_{ij} = A_j F_{ji}$. Indichiamo questi cinque fattori di forma con F_{TQ} , F_{TT} , F_{RT} , F_{QR} e F_{QQ} , rispettivamente. Osservando che le aree di questi elementi verificano

$$\begin{aligned}
A_Q &= 2 A_T \\
A_R &= 2^{1/2} A_Q
\end{aligned}$$

dalla relazione di reciprocità otteniamo

$$\begin{aligned}F_{QT} &= \frac{1}{2} F_{TQ} \\F_{TR} &= 2^{1/2} F_{RT} \\F_{QR} &= 2 F_{RQ}\end{aligned}$$

Ora applichiamo la relazione di conservazione dell'energia (ovvero dell'angolo solido), $\sum_j F_{ij} = 1$. Si ottiene il sistema lineare:

$$\begin{aligned}F_{QQ} + F_{QR} + 2 F_{QT} &= 1 \\F_{TT} + F_{TR} + 2 F_{TQ} &= 1 \\2 F_{RT} + 2 F_{RQ} &= 1\end{aligned}$$

Sostituendo in queste equazioni le precedenti identità ora abbiamo

$$F_{QQ} + 2 F_{RQ} + F_{TQ} = 1$$

$$F_{TT} + 2^{1/2} F_{RT} + 2 F_{TQ} = 1$$

$$2 F_{RT} + 2 F_{RQ} = 1$$

Per semplicità scriviamo $a = F_{QQ}$, $b = F_{TT}$, $c = F_{RQ}$, $d = F_{TQ}$, $e = F_{RT}$. Abbiamo tre equazioni in cinque incognite a , b , c , d , e i cui valori però sono vincolati ad appartenere all'intervallo $[0,1]$. Dall'ultima equazione ricaviamo $c = \frac{1}{2} - e$. Sostituendo nella prima otteniamo

$$a + (1 - 2e) + d = 1, \text{ ossia}$$

$$a + d = 2e.$$

La seconda equazione invece diventa $b + 2^{1/2}(\frac{1}{2} - e) + 2d = 1$, ossia

$$b + 2d - 2^{1/2}e = 1 - 1/2^{1/2}.$$

In tal modo restano due equazioni in quattro incognite. L'ipotesi della parte (a) del problema è $a=b=d=1/4$. Pertanto tre delle quattro incognite vengono fissate, e rimangono due equazioni in una sola incognita, che sono le seguenti:

$$\frac{1}{2} = a + d = 2e, \text{ ossia } e = 1/4,$$

e

$$b + 2d - 2^{1/2}e = 1 - 1/2^{1/2}, \text{ ossia } \frac{3}{4} - 2^{1/2}/4 = 1 - 1/2^{1/2}. \text{ ossia } 3 - 2^{1/2} = 4 - 2^{3/2}.$$

Quest'ultima identità è falsa, quindi i dati sono sufficienti ma incompatibili fra loro.

Nella parte (b) si richiede solo $a=d=1/4$. Pertanto restano due equazioni in due incognite, le seguenti: la prima è $e=1/4$, come prima, mentre la seconda diventa $b + 1/2 - 2^{1/2}/4 = 1 - 1/2^{1/2}$, ossia $b = 2^{1/2}/4 + 1/2 - 1/2^{1/2}$. Quindi si ottiene un valore di b dell'ordine di 0.15, ossia fra 0 e 1. I dati sono compatibili.

Esercizio 6. La scena consiste di una stanza cubica di lato 1 il cui pavimento è stato sostituito da una piramide il cui vertice alto tocca il centro del soffitto. Le quattro facce della piramide sono numerate da 1 a 4, le pareti da 5 a 8 (5 antistante a 1, 6 antistante a 2 e così via), ed il soffitto ha il numero 9. La piramide non riflette la luce, però la emette con potenza per unità di area uguale a 1. Le pareti 5 e 7 sono nere. Tutti gli altri coefficienti di riflettività sono $\frac{1}{2}$ tranne quello del soffitto che è $\frac{1}{4}$. I fattori di forma da ciascuna faccia della piramide alla parete antistante ed al soffitto valgono entrambi a , ed il fattore di forma da ciascuna parete al soffitto vale b , dove a e b sono due numeri fra 0 e 1.

1. Qual'è l'area di ciascuna faccia della piramide?
2. Se $a = b$, i dati sono sufficienti a trovare tutti i fattori di forma? Se sì, determinare i valori di a per cui i dati sono compatibili (ossia tali che tutti i fattori di forma siano compresi fra 0 e 1).
3. Se la risposta al punto precedente è negativa, assumiamo a diverso da b . Che condizioni dobbiamo imporre su a e b affinché i dati siano compatibili? Se si sono trovate ai punti precedenti valori di a e b per i quali i dati sono compatibili, si scelga a piacere una tale coppia a e b (uguali fra loro, se la domanda al punto (2) aveva risposta positiva), e si risponda alle seguenti domande:
4. Trovare le prime due iterazioni della radiosità ($b^{(1)}$, $b^{(2)}$) con il metodo di Jacobi.
5. Trovare $b^{(1)}$ con il metodo di Gauss-Seidel.
6. Trovare la seconda iterazione del metodo di Southwell.

Esercizio 7. Una stanza si compone di un cubo la cui faccia superiore è sostituita da un tetto fatto a forma di calotta a volta, con area doppia di quella delle facce del cubo. Il soffitto emette luce con potenza 1, il pavimento è nero, le pareti ed il soffitto hanno fattore di riflettività $\frac{1}{2}$. Il fattore di forma da una parete a quella opposta frontalmente vale α .

- i) La scelta di $\alpha = 1/5$ è compatibile con i dati del problema? Se sì prendere $\alpha = 1/5$, se no scegliere un altro valore di α compatibile.
- ii) Calcolare la prima iterazione della radiosità con il metodo di rilassamento di Jacobi.
- iii) Calcolare le prime due iterazioni del metodo di Jacobi.

Svolgimento.

Chiamiamo F_{lat} il fattore di forma da una faccia quadrata ad una faccia quadrata adiacente, e F_{front} quello da una faccia quadrata a quella frontale. La volta, vista dalle altre facce, copre lo stesso angolo solido che coprirebbe la faccia quadrata superiore del cubo, e quindi il fattore di forma da faccia laterale a volta vale ancora F_{lat} e quello da pavimento a volta vale F_{front} . Poiché la volta ha area doppia delle altre facce, dalla relazione di reciprocità $A_i F_{ij} = A_j F_{ji}$ otteniamo che il fattore di forma da volta a faccia laterale vale $\frac{1}{2} F_{lat}$ e quello da volta a pavimento vale $\frac{1}{2} F_{front}$. Infine, la volta è convessa invece che piana, e quindi ha un fattore di forma non nullo F_v verso se stessa. Allora la conservazione dell'energia, ossia il fatto che l'angolo solido totale sia il 100% (ovvero 1), porta alle seguenti identità:

$$\begin{aligned} 4 F_{lat} + F_{front} &= 1 && \text{(angolo solido visto dal pavimento o da ciascuna delle pareti)} \\ F_v + 2F_{lat} + \frac{1}{2} F_{front} &= 1 && \text{(angolo solido visto dalla volta)} \end{aligned}$$

Questo sistema lineare ha due equazioni e tre incognite, di cui una, F_{front} , è posta uguale ad un parametro fisso α compreso fra 0 e 1. Allora si può ricavare F_{lat} dalla prima

equazione e F_v dalla seconda, ma occorre verificare per quali valori di questi due risultati rimangono compresi fra 0 e 1. Si trova:

$$F_{lat} = (1 - \alpha)/4$$

sempre compreso fra 0 e 1 se α lo è, e

$$F_v = 1 - (1 - \alpha)/2 - \alpha/2 = \alpha$$

anch'esso compreso fra 0 e 1. Quindi per ogni valore ammissibile del parametro α si trova una soluzione.

Numeriamo con 1 il soffitto, con 2,3,4,5 le quattro pareti laterali e con 6 il pavimento. Poiché tutti i coefficienti di riflettività valgono $\frac{1}{2}$ tranne quello del pavimento che vale 0, la matrice M della radiosità è la seguente:

$$\begin{pmatrix} 1 - \frac{1}{2}\alpha & (\alpha - 1)/16 & (\alpha - 1)/16 & (\alpha - 1)/16 & (\alpha - 1)/16 & -\frac{1}{4}\alpha \\ (\alpha - 1)/8 & 1 & (\alpha - 1)/8 & -\frac{1}{2}\alpha & (\alpha - 1)/8 & (\alpha - 1)/8 \\ (\alpha - 1)/8 & (\alpha - 1)/8 & 1 & (\alpha - 1)/8 & -\frac{1}{2}\alpha & (\alpha - 1)/8 \\ (\alpha - 1)/8 & -\frac{1}{2}\alpha & (\alpha - 1)/8 & 1 & (\alpha - 1)/8 & (\alpha - 1)/8 \\ (\alpha - 1)/8 & (\alpha - 1)/8 & -\frac{1}{2}\alpha & (\alpha - 1)/8 & 1 & (\alpha - 1)/8 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Il vettore della potenza luminosa creata è $E = (1, 0, 0, 0, 0, 0)$, perché solo il soffitto emette luce. Il sistema lineare della radiosità, $Mb = e$, ci dà il vettore delle radiosità dei sei elementi, ma osserviamo che la radiosità del pavimento è 0 perché il pavimento ha coefficiente di riflettività 0 (essendo nero) e non emette luce propria. Quindi possiamo scartare l'ultima incognita e ridurci al seguente problema lineare a dimensione cinque:

$$\begin{pmatrix} 1 - \frac{1}{2}\alpha & (\alpha - 1)/16 & (\alpha - 1)/16 & (\alpha - 1)/16 & (\alpha - 1)/16 \\ (\alpha - 1)/8 & 1 & (\alpha - 1)/8 & -\frac{1}{2}\alpha & (\alpha - 1)/8 \\ (\alpha - 1)/8 & (\alpha - 1)/8 & 1 & (\alpha - 1)/8 & -\frac{1}{2}\alpha \\ (\alpha - 1)/8 & -\frac{1}{2}\alpha & (\alpha - 1)/8 & 1 & (\alpha - 1)/8 \\ (\alpha - 1)/8 & (\alpha - 1)/8 & -\frac{1}{2}\alpha & (\alpha - 1)/8 & 1 \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \\ b_5 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

Con abuso di notazione, continuiamo a denotare questo sistema lineare con $Mb = e$. Scriviamo $M = N - P$, con

$$N = \begin{pmatrix} 1 - \frac{1}{2}\alpha & 0 & 0 & 0 & 0 \\ (\alpha - 1)/8 & 1 & 0 & 0 & 0 \\ (\alpha - 1)/8 & (\alpha - 1)/8 & 1 & 0 & 0 \\ (\alpha - 1)/8 & -\frac{1}{2}\alpha & (\alpha - 1)/8 & 1 & 0 \\ (\alpha - 1)/8 & (\alpha - 1)/8 & -\frac{1}{2}\alpha & (\alpha - 1)/8 & 1 \end{pmatrix}$$

e

$$P = \begin{pmatrix} 0 & (1-\alpha)/16 & (1-\alpha)/16 & (1-\alpha)/16 & (1-\alpha)/16 \\ 0 & 0 & (1-\alpha)/8 & \frac{1}{2}\alpha & (1-\alpha)/8 \\ 0 & 0 & 0 & (1-\alpha)/8 & \frac{1}{2}\alpha \\ 0 & 0 & 0 & 0 & (1-\alpha)/8 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Il rilassamento di Jacobi approssima la soluzione **b** con le seguenti iterazioni:

$$b^{(k+1)} = N^{-1}Pb^{(k)} + N^{-1}e$$

Calcoliamo N^{-1} , ad esempio con il metodo di Gauss. Se si lascia arbitrario il valore di α , l'espressione di N^{-1} è assai complicata e laboriosa: la riportiamo qui di seguito per comodità del lettore. La matrice N^{-1} è triangolare inferiore con sulla diagonale i valori reciproci di quelli della matrice N : quindi

$$N_{11}^{-1} = \frac{1}{1-\alpha/2} = \frac{2}{\alpha-2}$$

e gli altri valori diagonali valgono 1. Per quanto riguarda i valori sotto la diagonale, si trova, riga per riga:

$$\begin{aligned} N_{21}^{-1} &= \frac{\alpha-1}{4(\alpha-2)} \\ N_{31}^{-1} &= \frac{9-10\alpha+\alpha^2}{64-32\alpha}; \quad N_{32}^{-1} = \frac{1-\alpha}{8} \\ N_{41}^{-1} &= \frac{-73+51\alpha+21\alpha^2+\alpha^3}{256(\alpha-2)}; \quad N_{42}^{-1} = \frac{1+30\alpha+\alpha^2}{64}; \quad N_{43}^{-1} = \frac{1-\alpha}{8} \\ N_{51}^{-1} &= \frac{649-476\alpha-226\alpha^2+52\alpha^3+\alpha^4}{2048(\alpha-2)}; \quad N_{52}^{-1} = \frac{65-3\alpha-61\alpha^2-\alpha^3}{512}; \quad N_{53}^{-1} = \frac{1+30\alpha+\alpha^2}{64}; \quad N_{54}^{-1} = \frac{1-\alpha}{8} \end{aligned}$$

Vediamo i risultati nel caso particolare $\alpha=1/5$, che è molto più agevole. Si trova

$$N^{-1} = \begin{pmatrix} 1.111111 & 0 & 0 & 0 & 0 \\ 0.111111 & 1 & 0 & 0 & 0 \\ 0.122222 & 0.1 & 1 & 0 & 0 \\ 0.134444 & 0.11 & 0.1 & 1 & 0 \\ 0.147889 & 0.121 & 0.11 & 0.1 & 1 \end{pmatrix}$$

e

$$P = \begin{pmatrix} 0 & 0.05 & 0.05 & 0.05 & 0.05 \\ 0 & 0 & 0.1 & 0.1 & 0.1 \\ 0 & 0 & 0 & 0.1 & 0.1 \\ 0 & 0 & 0 & 0 & 0.1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Cominciamo con l'approssimazione iniziale $\mathbf{b}^{(0)} = \mathbf{0}$. Si ottiene:

$$\mathbf{b}^{(1)} = \mathbf{N}^{-1}\mathbf{e} = \begin{pmatrix} 1.111111 \\ 0.111111 \\ 0.122222 \\ 0.134444 \\ 0.147889 \end{pmatrix}$$

Allora, i risultati sono:

$$\mathbf{b}^{(2)} = \mathbf{N}^{-1}\mathbf{P}\mathbf{b}^{(1)} + \mathbf{N}^{-1}\mathbf{e} = \mathbf{N}^{-1}\mathbf{P}\mathbf{b}^{(1)} + \mathbf{b}^{(1)} = \mathbf{N}^{-1} \begin{pmatrix} 0.0286481 \\ 0.0433204 \\ 0.0354302 \\ 0.0255288 \\ 0.0132927 \end{pmatrix} + \begin{pmatrix} 1.111111 \\ 0.111111 \\ 0.122222 \\ 0.134444 \\ 0.147889 \end{pmatrix} = \begin{pmatrix} 1.13976 \\ 0.154431 \\ 0.157652 \\ 0.159973 \\ 0.161182 \end{pmatrix}$$

Ovviamente, per simmetria, a parte il primo valore gli altri dovrebbero essere uguali, ma due sole iterazioni sono poche. Con tre iterazioni si trova

$$\mathbf{b}^{(3)} = \begin{pmatrix} 1.14629 \\ 0.16251 \\ 0.162996 \\ 0.163298 \\ 0.163509 \end{pmatrix}$$

e con quattro

$$\mathbf{b}^{(4)} = \begin{pmatrix} 1.14735 \\ 0.163715 \\ 0.163787 \\ 0.163836 \\ 0.163869 \end{pmatrix}$$

e con cinque

$$\mathbf{b}^{(5)} = \begin{pmatrix} 1.14751 \\ 0.1639 \\ 0.16392 \\ 0.163919 \\ 0.163924 \end{pmatrix}$$

Le iterazioni convergono a

$$b^{(\infty)} = \begin{pmatrix} 1.14754 \\ 0.163934 \\ 0.163934 \\ 0.163934 \\ 0.163934 \end{pmatrix}$$

Problema: quale è la differenza se si calcola la seconda iterazione del metodo di Gauss-Seidel? Spiegare le differenze passo per passo.

Esercizio 8. Una scena consiste dell'interno di un corridoio ad angolo, dato dall'unione di tre cubi di lato 1, il secondo ed il terzo dei quali si congiungono al primo su due pareti adiacenti di esso (ovviamente non ci sono pareti interne). La parete del terzo cubo opposta a quella dove esso si congiunge al primo cubo e' una finestra dalla quale entra luce a potenza per unita' di area pari a 1: ovviamente il suo coefficiente di riflettività e' zero. Le altre pareti non emettono luce; le pareti laterali ed il soffitto hanno coefficiente di riflettività $\frac{1}{2}$, il pavimento 0. Il pavimento, il soffitto e le sei pareti laterali, una delle quali e' la finestra, formano altrettanti elementi. Numeriamo il pavimento con 0, poi procedendo in verso orario (visto dall'alto) la finestra con 1, le due pareti corte che formano un angolo ottuso con 3 e 4, l'ultima parete corta con 5, le due pareti lunghe con 6 e 7, il soffitto con 8.

1. Quanti sono i fattori di forma indipendenti?
2. Supponiamo di assegnare valore α al fattore di forma dal pavimento al soffitto, β dalla parete lunga alla parete corta adiacente, γ dalla parete lunga alla parete corta ad essa parallela, δ da una parete lunga all'altra parete lunga, θ fra due pareti corte adiacenti che si vedono. Questi dati sono sufficienti a trovare tutti i fattori di forma? Spiegare.
3. Se lo sono, determinare gli altri fattori di forma, se no fissare il numero minimo indispensabile di altri fattori di forma necessario per poterli trovare tutti, e quindi trovare i restanti.
4. Attribuire ai fattori di forma valori numerici a piacere ma compatibili, e trovare le prime due iterazioni del metodo di Jacobi.

Esercizio 9. Come nell'[Esercizio 4 della sezione Esercizi sul calcolo dei fattori di forma](#), si consideri una ciotola K di forma emisferica di raggio 1 appoggiata nel suo punto più basso all'origine, ossia tangente all'origine al piano $\{z=0\}$ e contenuta nel semispazio superiore $\{z \geq 0\}$: quindi il bordo della ciotola è un cerchio J sul piano $\{z=1\}$. Poiché K è convessa, il fattore di forma dalla ciotola a sé stessa è positivo.

1. Si assuma che il fattore di forma F_{KK} sia 0.75, e che la ciotola abbia coefficiente di riflettività $\frac{1}{2}$. L'esterno della ciotola non riflette luce e si può approssimare quindi con uno sfondo emisferico nero (quindi la scena diventa l'interno di S). Al polo nord di S poniamo una luce puntiforme con potenza 1 (densità di energia per unità di tempo e area). Calcolare le prime quattro iterazioni del metodo di Jacobi per la radiosità.
2. Calcolare le prime tre iterazioni del metodo di Gauss-Seidel.
3. Calcolare le prime tre iterazioni del metodo di Southwell.

Svolgimento.

Parte 1. Sia C l'emisfero superiore nero. Ovviamente $F_{CC} = F_{KK} = 0.75$, e per la conservazione dell'angolo solido totale $F_{KC} = F_{CK} = 0.25$. Numeriamo l'elemento K con il numero 1 e l'elemento C con 2. Il vettore \mathbf{e} della potenza creata è $\mathbf{e}=(0,1)$. La matrice della radiosità è

$$M = \begin{pmatrix} 1 - \frac{1}{2} \cdot 0.75 & -\frac{1}{2} \cdot 0.25 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} \frac{5}{8} & -\frac{1}{8} \\ 0 & 1 \end{pmatrix}$$

Decomponiamo $M = N - P$, con

$$N = \begin{pmatrix} \frac{5}{8} & 0 \\ 0 & 1 \end{pmatrix}$$

$$P = \begin{pmatrix} 0 & \frac{1}{8} \\ 0 & 0 \end{pmatrix}$$

Chiaramente,

$$N^{-1} = \begin{pmatrix} \frac{8}{5} & 0 \\ 0 & 1 \end{pmatrix}$$

e quindi

$$N^{-1}P = \begin{pmatrix} \frac{8}{5} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & \frac{1}{8} \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & \frac{1}{5} \\ 0 & 0 \end{pmatrix}$$

Troviamo le prime iterazioni del *metodo di Jacobi*. Partiamo, come al solito, con l'approssimazione iniziale $\mathbf{b}^{(0)} = (0,0)$. Allora

$$\mathbf{b}^{(1)} = N^{-1}\mathbf{e} = (0,1)$$

(la seconda colonna di N^{-1}). Osserviamo che, per ogni t , si ha $N^{-1}P(t,1) = (1/5,0)$, e quindi

$$\mathbf{b}^{(2)} = N^{-1}P \mathbf{b}^{(1)} + N^{-1}\mathbf{e} = (1/5,0) + (0,1) = (1/5,1),$$

$$\mathbf{b}^{(3)} = N^{-1}P \mathbf{b}^{(2)} + N^{-1}\mathbf{e} = (1/5,0) + (0,1) = (1/5,1).$$

Pertanto, dopo il terzo passo, le iterazioni si stabilizzano: la radiosità della ciotola K è 1/5 della potenza emessa dalla lampada, il resto della luce si perde nello sfondo (ma naturalmente lo sfondo continua a creare luce con densità di potenza pari a 1).

Parte 2. Ora rieseguiamo il calcolo delle prime tre iterazioni con il *metodo di Gauss-Seidel*. Nel calcolo di $\mathbf{b}^{(1)} = N^{-1}\mathbf{e}$, la prima componente del risultato è 0, ossia non cambia

rispetto a $\mathbf{b}^{(0)}$, e pertanto anche nel calcolo della seconda componente la prima iterazione di Gauss-Seidel coincide con quella di Jacobi:

$$\mathbf{b}^{(1)} = (0, 1).$$

Nella seconda iterazione, applichiamo la prima riga della matrice $N^{-1}P$ a $\mathbf{b}^{(1)}$, ottenendo così la prima componente b_{21} di $\mathbf{b}^{(1)}$, che, come prima, vale $1/5$. Invece la seconda riga, per il calcolo della seconda componente b_{22} di $\mathbf{b}^{(2)}$, viene applicata al vettore che ha come prima componente quella appena ottenuta. Ma la seconda riga di $N^{-1}P$ è nulla, e quindi non cambia niente: i due metodi coincidono alla seconda iterazione. Per lo stesso motivo, si vede che al passo successivo metodi continuano a coincidere, ed a quel punto le iterazioni si sono stabilizzate.

Parte 3. Ora rieseguiamo il calcolo col *metodo di Southwell*. Alla n -sima iterazione, dobbiamo trovare quale è la componente più grande del vettore resto $\mathbf{r}^{(n)} = \mathbf{e} - M \mathbf{b}^{(n)}$ e rilassare solo quella. Per $n=0$ ovviamente si ha $\mathbf{r} = \mathbf{e}$, e quindi rilassiamo la seconda riga (la prima componente di \mathbf{e} è 0). Il risultato è $\mathbf{b}^{(1)} = (0, 1)$ (ossia la prima componente non cambia): questo era anche il risultato degli altri metodi, i quali però richiedevano di risolvere entrambe le righe del sistema lineare e non una sola (è vero che l'altra è banale, ma il software di esecuzione questo non può saperlo e quindi non può tenerne conto). Ora abbiamo $\mathbf{r}^{(1)} = \mathbf{e} - M \mathbf{b}^{(1)} = (0, 1) - (-1/8, 1) = (1/8, 0)$. La componente più grande è la prima, e quindi rilassiamo la prima riga. Come già visto, si ottiene come prima componente di $\mathbf{b}^{(2)} = N^{-1}P \mathbf{b}^{(1)} + N^{-1}\mathbf{e}$ il valore $1/5$. La seconda componente non cambia, resta quindi 1, ed il risultato è $\mathbf{b}^{(2)} = (1/5, 1)$: lo stesso risultato degli altri due metodi. Allo stesso modo, si trova $\mathbf{b}^{(3)} = (1/5, 1)$ ed a questo punto l'iterazione si è stabilizzata: ma abbiamo eseguito solo la metà delle operazioni aritmetiche (in effetti, abbiamo evitato di risolvere quelle equazioni lineari banali i cui coefficienti sono tutti nulli).